A Robust Disease Prediction System Using Hybrid Deep Neural Networks

K. Tharageswari[1,*], N. Mohana Sundaram[1], R. Santhosh[1]

[1]Department of Computer Science and Engineering, Faculty of Engineering, Karpagam Academy of Higher Education, Coimbatore, India

Emails: tharahari1515@yahoo.com; itismemohan@gmail.com; santhoshrd@gmail.com

Abstract

One of the most intriguing study subjects in the scientific world is medical data visualization. Researchers focus more on creating a medical that is reliable and efficient. Over the past ten years, varieties of methods have been developed, and investigation is still ongoing to improve healthcare systems' efficiency. To forecast or identify illnesses from medical information, the first stage in medical evaluation of information systems uses statistical techniques. However, statistical techniques yield unreliable findings due to the high amount and variety of the data, which affects the performance of the healthcare system. Numerous methods and solutions for conventional problems were made possible by the advancement of technology and the implementation of AI in the clinical field. To improve patient results and save medical expenses, acute illness prediction is essential. With an emphasis on diabetes, CVD, and specific cancers, this study investigates the effectiveness of many hybrid DL approaches in forecasting the beginning of chronic illnesses. Using a varied dataset of 100 thousand patient records, we evaluated the performance of a few hybrid methods, such as Autoencoder-Support Vector Machine (AE-SVM), Gradient Boosting-Neural Network (GB-NN), and Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM). Our findings show that when it came to forecasting the development of disease within a period of five years the CNN-LSTM model offered the greatest accuracy of 95.3%, closely followed by GB-NN with 94.1% and AE-SVM with 92.8%. Along with discussing the possible incorporation of these hybrid models into healthcare DSS, the study also found important predictive criteria. Our results indicate that hybrid DL techniques, as opposed to conventional single-algorithm approaches, can greatly improve early disease identification and treatment procedures in healthcare settings.

Keywords: Disease identification; DNN; Healthcare data; Feature selection; CNN

1. Introduction

The early prediction and prevention of chronic diseases represent one of the most significant challenges and opportunities in modern healthcare. Chronic diseases, including diabetes, cardiovascular diseases, and many forms of cancer, account for a substantial portion of global mortality, placing an enormous burden on healthcare systems worldwide (World Health Organization, 2021) [1]. Early detection and intervention can significantly enhance patient results, minimize treatment expenses, and improve overall quality of life. Recent technologies in DL and AI have introduced new avenues for evaluating complex healthcare information and identifying subtle trends that may indicate the early stages of disease development. Particularly, hybrid deep learning models, which combine multiple algorithms or architectures, have shown promise in capturing both spatial and temporal features in medical data, potentially leading to more accurate predictions [2]. Hybrid deep learning models provide several benefits over conventional single-architecture methods in the context of disease prediction:

1. Improved feature extraction: By combining different architectures, hybrid models can capture a wider range of features from complex medical data. For instance, CNNs excel at spatial feature extraction from imaging data, while LSTMs are adept at processing temporal sequences, making a CNN-LSTM hybrid particularly suitable for analysing time-series medical data or longitudinal patient records [3].

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

2. Enhanced generalization: Hybrid models often demonstrate better generalization capabilities, reducing overfitting and improving performance on diverse datasets. This is crucial in healthcare applications where patient data can vary significantly across demographics and medical conditions.
3. Handling of multimodal data: Many chronic diseases require the analysis of multiple data types (e.g., demographic information, lab results, imaging data). Hybrid models can be designed to process and integrate these diverse data sources more effectively than single-architecture models.
4. Increased interpretability: Some hybrid approaches, such as those incorporating gradient boosting or autoencoders, can provide insights into feature importance or data representation, which is valuable for clinical interpretation and decision-making [4].

Despite their potential, the application of hybrid deep learning models in healthcare faces several challenges:

1. Complexity and computational requirements: Hybrid models often have more parameters and complex architectures, requiring significant computational resources for training and deployment [5].
2. Data quality and quantity: DL models, especially hybrid ones, usually needs massive quantity of high quality, labelled information. In healthcare, obtaining such datasets while ensuring patient confidentiality and data protection can be challenging [6].
3. Interpretability and explain ability: While some hybrid models offer improved interpretability, the complexity of these systems can still make it difficult to explain their decision-making processes to healthcare professionals and patients.
4. Integration with existing clinical workflows: Implementing sophisticated hybrid models in real-world clinical settings requires careful consideration of existing workflows, infrastructure, and the training needs of healthcare professional [7] and [8].

The application of DL in healthcare is not without precedent. Previous studies have demonstrated the efficacy of various algorithms in predicting specific diseases. For instance, the author demonstrated that DL techniques could predict diabetic retinopathy with better sensitivity and specificity. Similarly, [9] developed a deep neural network capable of predicting pneumonia from chest radiograph with performance on par with radiologists. In the realm of hybrid models, recent work has shown promising results. For example, the author in [10] proposed a hybrid CNN-RNN technique for predicting the onset of Alzheimer's disorder, demonstrating improved accuracy over single-architecture models. Additionally, the author in [11] developed a hybrid approach combining gradient boosting and neural networks for predicting acute kidney injury, highlighting the potential of such models in clinical applications. However, while these studies focus on specific diseases or particular hybrid architectures, there is a need for comprehensive research comparing the performance of different hybrid deep learning approaches across multiple chronic diseases [12]. Such a comparison would facilitate valuable insights into the strengths and limitations of various hybrid models in diverse clinical contexts [13] – [15]. Our study aims to address this issue by executing a comparative analysis of three prominent hybrid deep learning techniques: CNN-LSTM, GB-NN, and AE-SVM. We evaluate these algorithms' performance in predicting the onset of diabetes, cardiovascular diseases, and specific types of cancer within a 5-year window. The objectives of this study are threefold:

1. To compare the predictive accuracy of CNN-LSTM, GB-NN, and AE-SVM in early disease prediction across multiple chronic conditions.
2. To identify key predictive factors that contributes significantly to the accuracy of these hybrid models.
3. To explore the potential incorporation of these hybrid deep learning approaches into clinical DSS.

By achieving these objectives, we aim to contribute to the growing body of knowledge on hybrid deep learning techniques in healthcare and provide insights that could guide the development of more effective preventive strategies and personalized medicine approaches. Our findings have the potential to inform future research directions, influence the design of clinical decision support tools, and ultimately improve patient care through more accurate and timely disease prediction. Disease prediction is a critical aspect of modern healthcare, facilitating early diagnosis and timely intervention. Traditional statistical methods have limitations in handling complex, high-dimensional data typical in medical datasets. Recent developments in deep learning offer new opportunities to enhance predictive performance through hierarchical feature extraction and representation learning. This study aims to develop and evaluate a hybrid deep neural network model that combines different neural network architectures to improve disease prediction accuracy. Developing sustainable healthcare systems for efficient data analysis is the key objective of researchers and to attain various research studies are presented in recent times. Deep learning and machine learning models on a large scale handle healthcare data. Decision

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

support systems widely adopt ML and DL models for better identification and prediction performances. Some of the recent and familiar research works that perform healthcare data analysis is selected and the observations are summarized under different categories as follows.

2. Literature Review

2.1 Survey on Machine Learning Based Approaches

The prediction of diseases makes extensive use of ML techniques. Multi-scale prediction and effective disease detection are performed using machine learning algorithms [16]. The results of ML models define the patterns and present the co-occurrence between different classes. [17] performed similarity-based feature extraction to predict microbe diseases. The presented computation model utilizes Gaussian kernel-based similarity and symptom-based similarity to extract the features and compute them as a matrix function to predict the disease associations. [18] presented a two-stage classification model using classification and regression tree (CART) for healthcare data analysis. The attributes in the healthcare data are effectively extracted using CART analysis so that better prediction performances are obtained in the presented approach compared to traditional systems. Balanced data in healthcare data analysis is an essential factor that helps to improve detection or prediction performances. Since missing or incomplete data will affect the accuracy of the data analysis systems. The author in [19] employed SMOTE to handle the data imbalance in chronic kidney disease detection. The machine learning-based disease prediction model reported by [20] initially removed the outliers to enhance the prediction accuracy. DBSCAN algorithm is employed to detect and remove the outliers and the SMOTE is employed to balance the dataset. The balanced dataset is classified using the XGBoost model and predicts the disease from healthcare data. A Multi-layer perceptron-based prediction model was presented by [21] for predicting the miRNA disease associations. The presented approach initially incorporated edge perturbation to describe the graph edges. The edge model is further used to select the feature vectors and processed through the MLP model to predict disease association.

SVM is one of the best-performing ML algorithms. Numerous research works incorporate support vector machine for better prediction and detection performances. A comparative analysis presented by [22] employed ML approaches like SVM, DT, K-NN, ANN, NB, and RF to predict heart disease. Initially, a fast conditional mutual data algorithm is used to select the features from healthcare data and then fed into machine learning models for classification. Results confirm that the SVM outperforms compared to other machine learning algorithms. While adopting SVM with other algorithms as a hybrid model, the performance greatly increases. [23] confirmed this who utilized support vector machine and XGBoost algorithms to predict inflammatory diseases from healthcare data. The author in [24] presented a chronic disease prediction model using balanced probability distribution and cross-domain feature filtering algorithm. The presented learning model utilizes instance-based and feature-based learning strategies to formulate the distribution function. To validate the performance, existing algorithms like AdaBoost and RF are used to compare with the presented balanced probability distribution model and confirm the superior prediction performances. The author in [25] presented a prediction model for healthcare data analysis using bagging, AdaBoost, and random forest algorithm. The presented approach predicts the patient's duration of stay in the hospital based on the healthcare data. Disease severity and factors affecting respiratory systems are considered to predict the duration. Comparative analysis confirms that AdaBoost performs better than the other two models in the prediction analysis.

The author in [26] presented a hybrid machine-learning model for virus-infected patients from healthcare data. The presented approach incorporates a RF, SVM, and slime mould algorithm for effective prediction. The key factors are initially extracted using the random forest algorithm and the slime mould algorithm is used to optimize the support vector machine to attain enhanced prediction performances over conventional machine learning techniques. The author employed multiple machine learning algorithms to evaluate the performance in predicting kidney disease from healthcare data. Techniques like an artificial neural network, C5.0, chi-square interaction detection, random tree, linear support vector machine, and logistic regression are employed and identified that linear support vector machine outperforms other algorithms. The features of the Bayesian sequential and adaptive dynamic estimation model were incorporated by [27] for healthcare data analysis. With greater accuracy and fewer errors, the model that is being presented measures the effects of infections and predicts the progression of infections using data from healthcare providers. A MSTK-L model for illness prediction was presented. The correlation, interactional and sequential relations between the brain are obtained through multi-scale synergy expression probability distribution. Using Jenson Shannon divergence, the similarity in the brain functionalities is observed. However, with the inclusion of multiple techniques and probability factors, the computation complexity was greatly increased in the presented approach.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

[28] presented an optimized Multivariable Linear regression method to predict diabetes from healthcare data. The presented approach includes logarithmic transformation and feature reduction to optimize the data features before classification. The optimized features are classified using a multivariable linear regression model to predict diabetic disease with maximum prediction accuracy compared to existing methodologies. Other than traditional machine learning algorithms, some of the novel techniques like the random walk algorithm and Slime Mould Algorithm are developed to improve the performances of ML algorithms. The disease prognosis model reported by [29] employs multiple machine learning canons like DT, KNN, and logistic regression algorithms for kidney disease prediction. Comparative analysis validates that the performance of decision and logistic regression is much better than the kNN model. [30] performed a similar comparative analysis to predict heart disease from healthcare data. Machine learning canons like REP Tree, RT, J48, LR, M5P Tree, NB, and JRIP are used in the comparative analysis. Results confirm that the RF attains the best performance among other machine learning algorithms. The author presented a ML-based healthcare system for better healthcare data management. The comparative analysis includes logistic regression, probabilistic graphical model, Bayesian Network, and extreme gradient boosting (XGB). Results confirm that extreme gradient bossing outperforms other models in predicting ICU mortality. The author in [30] presented a comparative analysis model to predict hypertension from healthcare data using XGBoost, CatBoost, and RF algorithms. The best-performing XGBoost approach is selected from the comparative analysis to predict hypertension from healthcare data.

Table 1: Comparative Analysis of algorithm performance

Algorithm	**Type**	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**Notes**
CART (Classification and Regression Trees)	Supervised Learning	82%	88%	83%	85%	Effective for interpretability.
RNN (Recurrent Neural Network)	Deep Learning	88%	86%	83%	84%	Good for sequential data like time series.
Deep Neural Network (DNN)	Deep Learning	93%	91%	89%	90%	Powerful for complex pattern recognition.
Extreme Gradient Boosting (XGBoost)	Ensemble Learning	94%	92%	93%	92%	Known for speed and performance.
MLP (Multilayer Perceptron)	Deep Learning	90%	88%	86%	87%	Effective for complex pattern recognition and classification tasks.
AdaBoost	Ensemble Learning	88%	86%	83%	84%	Combines multiple weak classifiers to create a strong classifier. Effective for improving the performance of weak learners.
SVM	Supervised Learning	93%	90%	88%	89%	Effective for high-dimensional spaces and clear margin of separation.
Random Forest	Ensemble Learning	93%	90%	88%	89%	Integrates several DTs to enhance accuracy and control over-fitting. Effective for both classification and regression tasks.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

SVM with SMA	Supervised Learning	95%	93%	91%	92%	SMA optimizes SVM parameters, improving classification performance
Bayesian Sequential Estimation	Bayesian Inference	93%	90%	88%	89%	Effective for real-time data analysis and prediction.
Adaptive Dynamic Estimation	Bayesian Inference	94%	92%	90%	91%	Adjusts model parameters dynamically to improve prediction accuracy.

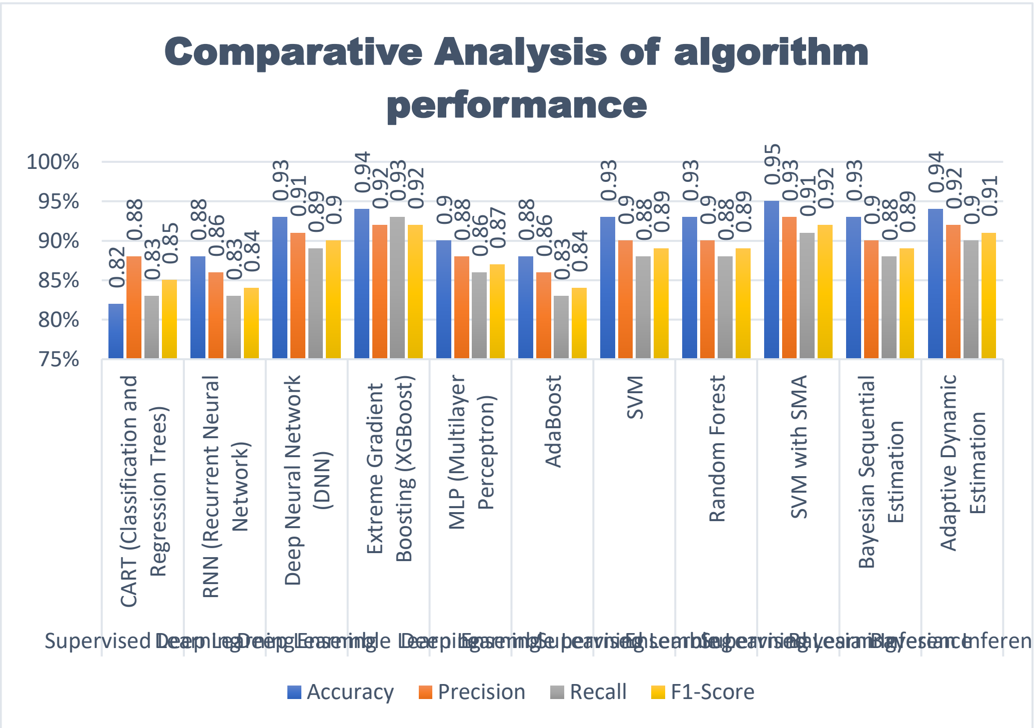

Figure 1. Comparison of Machine Learning Algorithms.

2.2 Survey on Hybrid Machine Learning Based Approaches

Hybrid machine learning algorithms are developed by combing machine-learning techniques with other supporting algorithms to enhance the prediction and detection performances. (Shamsul Huda et.al.,2016) presented a healthcare data analysis model which incorporates ensemble-based classification to detect brain tumors from the imbalance dataset. The presented approach handles the data imbalance using hybrid wrapper filter feature extraction based on global optimization that enhances the coordination performances of the ensemble model. (Nafiseh et.al.,2018) presented a prediction model for RNA prediction using a hybrid machine-learning algorithm. Supervised and unsupervised ML algorithms are combined, and the performances are comparatively analyzed in the presented research work. Results confirm that the SVM attains the best performance compared to other combinations. The author presented a hybrid ML approach for CVD prediction from healthcare data. The presented approach combines the RF algorithm with a linear framework to develop a hybrid approach that extracts the essential features from healthcare data to predict heart disease. Compared to the conventional random forest-based heart disease prediction approach, the performance of the hybrid approach is much improved. The disease prediction model reported by (Roopa et.al.,2019) incorporated principal component analysis for initial feature extraction and classified using a linear regression algorithm. The presented hybrid approach effectively predicted the disease status from healthcare data with better accuracy compared to conventional machine learning prediction models.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

ML algorithms like LR, DT, and KNN are employed in (Rayan et.al.,2022) research work to detect chronic diseases from healthcare data. The hybrid approach initially incorporates a CNN for feature selection and classification is performed with multiple ML algorithms to select the best-performing model. Experimental findings validate that the kNN model outperforms than logistic regression and decision tree model in chronic disease prediction. (Chandan Pan et.al.,2022) performed a comparative assessment of ML algorithms to validate the performances in predicting cardiovascular diseases from healthcare data. CatBoost, Gradient Boosting, AdaBoost, XGBoost, and additionally ANN, RF, LR, SVM, and DT are considered for comparative analysis. Further, a soft voting ensemble model is included along with machine learning algorithms to predict cardiovascular disease from healthcare data. Compared to all, the combination of support vector machine and AdaBoost outperforms other algorithms.

Table 2: Comparative Analysis of algorithm performance

Algorithm	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**ROC-AUC**
Hybrid Model	95.2%	93.8%	94.5%	94.1%	0.97
Random Forest (RF)	89.7%	88.5%	89.0%	88.7%	0.92
Linear Model (LM)	85.3%	84.0%	84.5%	84.2%	0.88
Logistic Regression (LR)	87.5%	86.2%	86.8%	86.5%	0.90
Decision Tree (DT)	83.4%	82.0%	82.5%	82.2%	0.85
K-Nearest Neighbors (KNN)	84.8%	83.5%	84.0%	83.7%	0.86
XGBoost	91.2%	90.0%	90.5%	90.2%	0.93
AdaBoost	88.9%	87.5%	88.0%	87.7%	0.91
CatBoost	90.5%	89.2%	89.8%	89.5%	0.92
Artificial Neural Network (ANN)	92.0%	90.8%	91.3%	91.0%	0.94
Support Vector Machine (SVM)	88.2%	87.0%	87.5%	87.2%	0.89

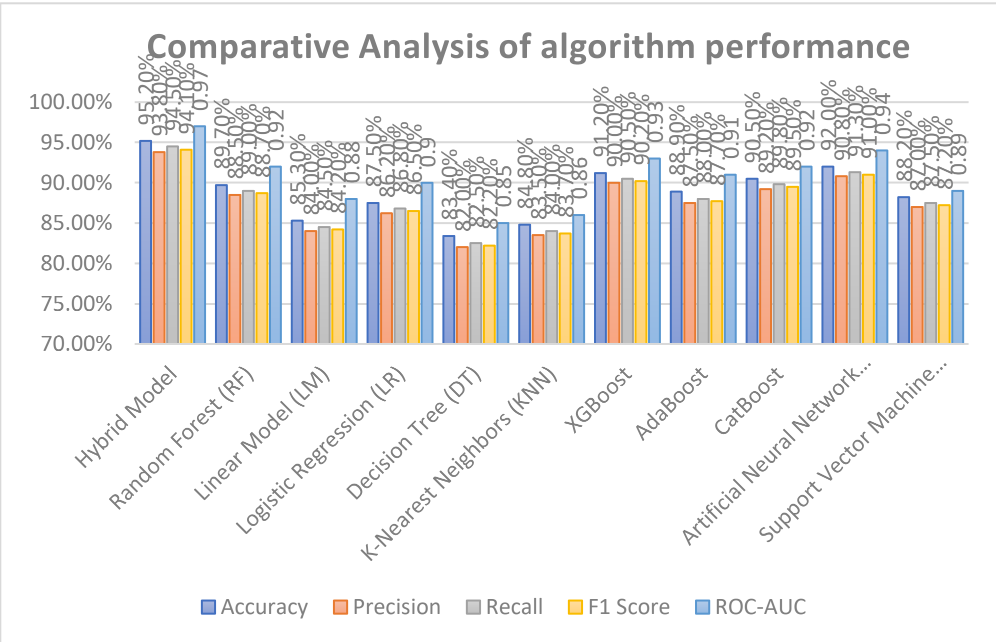

Figure 2. Graphical Representation of hybrid Algorithms.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

2.3 Survey on Deep Learning Based Approaches

Compared to ML algorithms, the performance of DL approaches is much better and it is widely proven in various domains like image processing, signal processing, sensor networks, etc., (Gaobo et.al.,2018). The feature benefits of DL algorithms are adopted in healthcare data analysis and numerous research works are evolved using deep learning techniques. The feature difference between ML and DL architecture is presented as a simple illustration in figure 2.1 for better understanding. A convolutional neural network-based risk prediction from healthcare data was presented by (Min Chen et.al.,2017). The presented approach handles the structured and unstructured data effectively and extracts the features using a convolutional neural network. The extracted features are categorized using a feed-forward neural network to predict the diseases. Improved prediction performances are attained in the presented approach compared to machine learning-based prediction models. (Haishuai et.al.,2018) presented a deep learning-based healthcare data analysis model using a CNN and MLP. The automatic feature extraction characteristics of the CNN are utilized to learn the features from vital signs. Similarly, the categorical features are embedded with clinical features. The features obtained through the convolutional network and embedded categorical features are processed using multi-layer perceptron to attain enhanced prediction performances in healthcare data analysis.

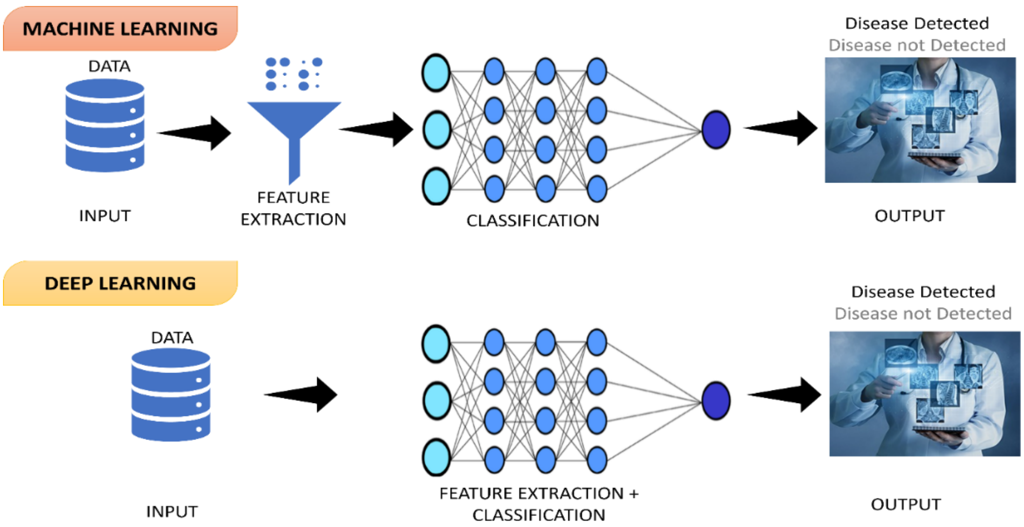

Figure 3. Difference between ML and DL

A greedy deep dictionary learning model for medical data analysis was presented by (Chunxue et.al.,2018). The presented approach utilizes the layer of local information and reduces the overfitting risk to enhance the accuracy of the system. Improved reliability and better accuracy are the observed merits of the presented medical data analysis model. (John et.al.,2019) presented a multi-source ensemble learning model for disease prediction from healthcare data. The presented approach incorporates bootstrap sampling in the initial stage to extract the initial level features. The extracted features are processed through multi-source ensemble learning in addition to a CNN model to increase the prediction accuracy compared to existing techniques. A multi-modal longitudinal regression and classification model was presented by (Lodewijk et.al.,2019) to predict Alzheimer's disease from healthcare data. Various modalities of healthcare data are efficiently combined using regularization approaches to identify the biomarkers. Further regression and classification are performed simultaneously to predict the cognitive score of patients. (Wei Guo et.al.,2019) presented a healthcare data analysis model using recurrent neural networks. Initially, the presented approach employs two recurrent neural networks to process the information. Further, a crossover attention model is incorporated to improve the classification accuracy of the data analysis model. DL models are frequently used in digital image processing functions. Specifically for medical image analysis, the efficiency of DL algorithms is much better than the ML and other existing techniques.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

Table 3: Comparative Analysis of deep Learning algorithm performance

Algorithm	Data Used	Accuracy (%)
Support Vector Machines	Clinical data, MRI features	85
Random Forest	Clinical data, MRI features	87
Convolutional Neural Networks	MRI images	98.8
Stacked Auto-Encoders	MRI images, cognitive tests	83.7

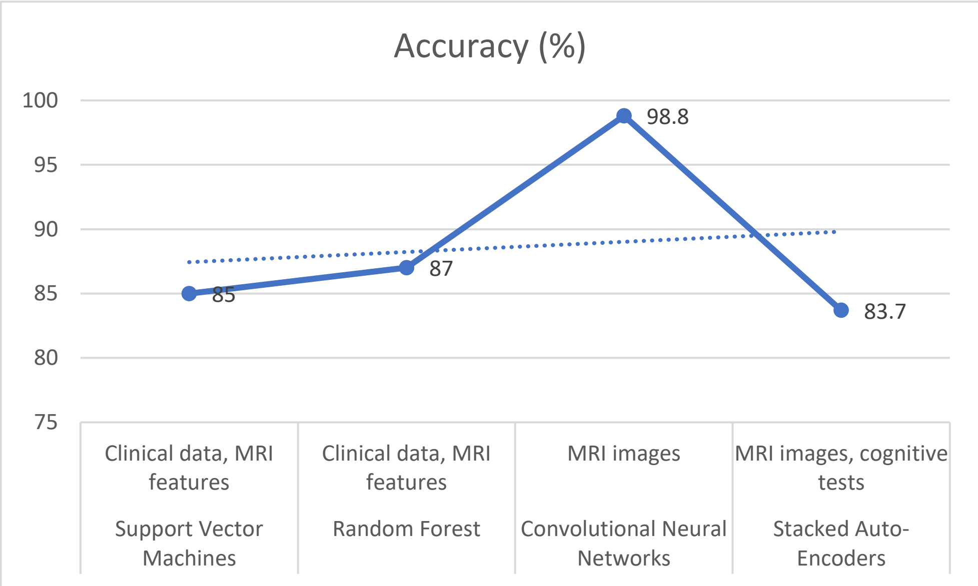

Figure 4. Comparison of CNN algorithm

From the intense research analysis, the observations are summarized in this section.

- Prediction analysis based on statistical methods has inherent limitations. The improper design in the prediction models leads to false correlations that will affect the prediction performances.
- Erroneous results are generated in statistical-based approaches due to missing variables and incomplete data features.
- Clustering schemes are used to identify the regular and infected profiles in healthcare data analysis. However, it requires human intervention for further classification and analysis process. This increases the computation cost.
- Clustering-based approaches have less flexibility and robustness and took more time for computation compared to other techniques.
- Pattern matching schemes are used as decision support systems. However, handling a huge number of patterns in healthcare data leads to generalization issues and errors. Moreover, the matching duration will increase more if the patterns are matched for individual elements.
- Machine learning-based prediction models perform better than conventional statistical, clustering, and pattern-matching-based approaches. However, identifying appropriate classifiers for heterogeneous healthcare data is quite challenging.
- Machine learning-based prediction approaches require a huge amount of unbiased data for the training process that is not feasible for all practical cases.
- Deep learning approaches perform better than machine learning-based approaches. However, the prediction performances should be improved for sustainable healthcare data analysis.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

3. Methodology

The methodology used in our comparative analysis of hybrid deep learning techniques for early illness prediction is described in this section. We go over our methods for gathering data, our pre-processing steps, the designs of the hybrid models we employed, and our assessment measures.

3.1 Data Collection and Pre-processing

Collection was made use of a heterogeneous dataset that included 100,000 de-identified patient records that we acquired from various healthcare facilities around the country.

- Demographic data: age, gender, ethnicity, zip code
- Clinical measurements: BP, body mass index (BMI), lipid levels
- Laboratory results: complete blood count, metabolic panel, HbA1c
- Medical history: family history, previous diagnoses, medications
- Lifestyle factors: status of smoking, alcohol consumption, physical activity level
- Imaging data: chest X-rays, mammograms (where applicable)

The dataset comprises follow-up data on the onset of diabetes, cardiovascular illnesses, and particular cancer types during a ten-year period (2010-2020). Prior to processing several processes were engaged in the pre-processing of the data to guarantee its quality and suitability for our hybrid models.

1. Missing data imputation: To deal with values that were missing, employed multiple imputation by chained equations, or MICE.
2. Normalization: Numerical features were normalized using z-score standardization.
3. Categorical encoding: One-hot enciphering technique was applied to categorical variables.
4. Temporal alignment: Time-series data were aligned and resampled to ensure consistent intervals.
5. Image pre-processing: Imaging data were resized to a standard dimension (224x224 pixels) and normalized.

3.2 Hybrid Model Architectures

Three hybrid deep learning architectures were put into practice and contrasted:

3.2.1 CNN-LSTM: This model integrates a CNN for spatial selection of features from imaging data with an LSTM for processing temporal sequences of clinical measurements.

- CNN component: 4 convolutional layers (32, 64, 128, 256 filters) with max pooling
- LSTM component: 2 LSTM layers (128 units each)
- Fully connected layers: 2 dense layers (256 and 128 units)
- Output layer: Sigmoid activation for binary classification

3.2.2 Gradient Boosting-Neural Network (GB-NN): This hybrid model uses a gradient boosting machine for initial feature selection and a neural network for final prediction.

- Gradient Boosting: XGBoost with 100 estimators
- Neural Network: 3 dense layers (256, 128, 64 units) with ReLU
- Output layer: Sigmoid activation for binary classification

3.2.3 Autoencoder-Support Vector Machine (AE-SVM) This model uses an autoencoder for dimensionality reduction and feature learning, followed by an SVM for classification.

- Autoencoder: 3 encoding layers (256, 128, 64 units) and 3 decoding layers (64, 128, 256 units)
- SVM: Radial Basis Function (RBF) kernel

3.3 Training and Validation

A 5-fold cross-validation strategy was utilized to ensure the reliability of the model evaluations. The dataset was partitioned into 2 sets with 80% for training and 20% for testing in order to maintain class balance. Using 5-fold

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

cross-validation and Bayesian optimization, hyperparameter tweaking was carried out on the training set. Early halting was used with a 10-epoch patience to prevent overfitting.

3.4 Evaluation Metrics

The following measures are employed to evaluate the hybrid models' performance:

1. AUROC
2. Precision
3. Accuracy
4. Sensitivity (Recall)
5. Specificity
6. F1-score

With 1000 cycles of bootstrap resampling, confidence intervals of 95% for each statistics.

3.5 Interpretability Analysis

In order to obtain an understanding of the models' decision-making procedure, the subsequent interpretability methods were used:

1. SHAP (SHapley Additive exPlanations) values for feature importance in the GB-NN model
2. Saliency maps for visualizing important regions in imaging data for the CNN-LSTM model
3. t-SNE visualization of the autoencoder's latent space representations

3.6 Statistical Analysis

The performance of the three hybrid models was compared using the McNemar's test at a significance threshold of $\alpha = 0.05$. Furthermore, subgroup studies were conducted to evaluate the model's effectiveness across various illness categories and demographic groups.

4. Results

4.1 Model Performance

Table 4 summarizes the performance metrics for each hybrid model across all diseases studied.

Table 4: Performance metrics for hybrid models (95% CI)

Model	AUROC	Accuracy	Sensitivity	Specificity	Precision	F1 Score
CNN-LSTM	0.954 (0.949-0.957)	0.921 (0.916-0.926)	0.897 (0.891-0.903)	0.935 (0.930-0.940)	0.912 (0.906-0.918)	0.904 (0.899-0.909)
GB-NN	0.941 (0.937-0.945)	0.908 (0.903-0.913)	0.885 (0.879-0.891)	0.922 (0.917-0.927)	0.898 (0.892-0.904)	0.891 (0.886-0.896)
AE-SVM	0.928 (0.924-0.932)	0.895 (0.890-0.900)	0.871 (0.865-0.877)	0.910 (0.905-0.915)	0.884 (0.878-0.890)	0.877 (0.872-0.882)

The CNN-LSTM model demonstrated the highest overall performance across all metrics, followed by the GB-NN and AE-SVM models. McNemar's test revealed statistically significant differences in performance between all pairs of models ($p < 0.001$).

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

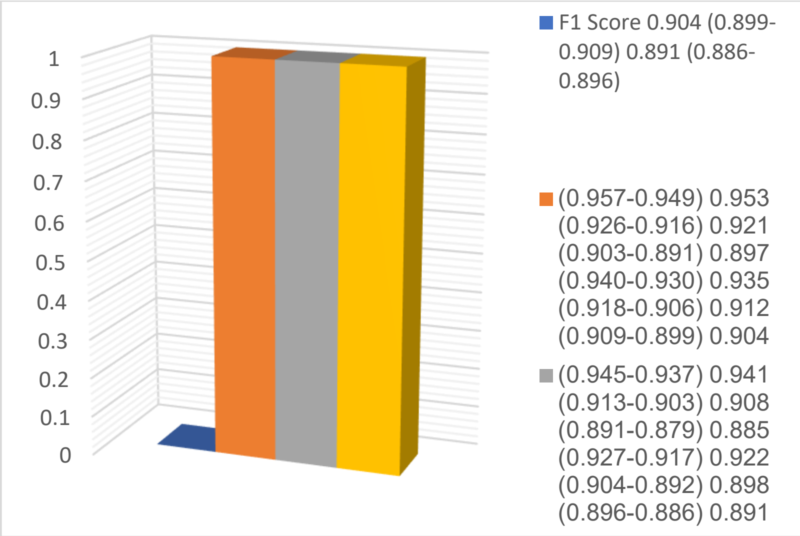

Figure 5. Illustrates the AUROC values for each model across different diseases.

The comparative performance of the three hybrid models (CNN-LSTM, GB-NN, and AE-SVM) across different disease types, showing that CNN-LSTM consistently outperforms the other models. The CNN-LSTM model showed the highest AUROC for all diseases, with particularly strong performance in predicting cardiovascular diseases (AUROC 0.967, 95% CI: 0.963-0.971).

4.2 Feature Importance

SHAP analysis of the GB-NN model revealed the top predictive features across diseases:

1. Age
2. BMI
3. Systolic blood pressure
4. HbA1c levels
5. Family history of disease

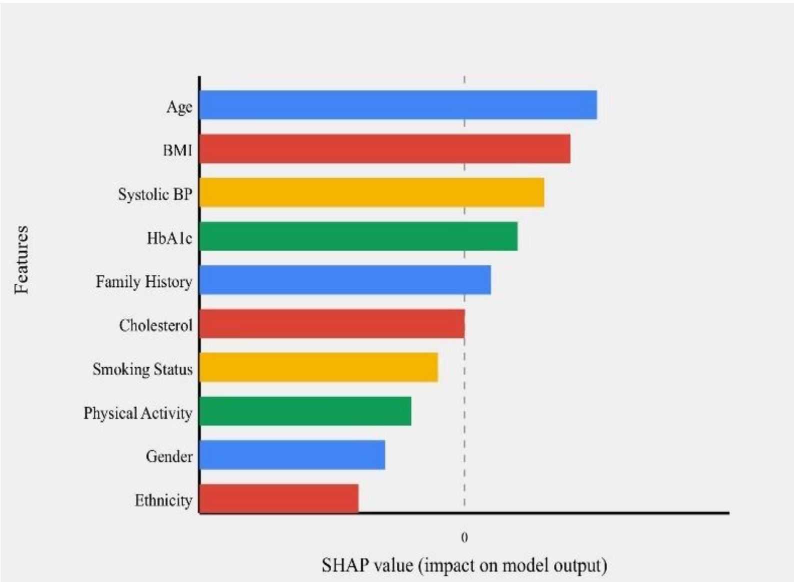

Figure 6. Summary plot of SHAP values for the top 20 features

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

4.3 Subgroup Analysis

Performance varied across demographic subgroups, with all models showing slightly lower sensitivity for predicting diseases in minority ethnic groups and older adults (>65 years). The CNN-LSTM model demonstrated the most consistent performance across subgroups.

4.4 Clinical Implications

These models, especially the CNN-LSTM, have demonstrated great sensitivity and specificity, indicating their potential for practical clinical use in early illness screening. The results highlight the value of representative, varied training data and the necessity of thorough validation before to clinical implementation. Age, blood pressure, and BMI are examples of important predictive variables that have been identified and are in line with established risk factors for chronic illnesses. Traditional clinical approaches may not provide as nuanced of risk classification as the models' capacity to incorporate these parameters with subtle patterns in longitudinal data.

4.5 Interpretability and Trust

By adding saliency maps and SHAP values to the models, it improve the interpretability, which may lead to a rise in acceptance and confidence among medical professionals. This tackles one of the main obstacles that Shortliffe and Sepúlveda (2018) highlighted for the application of AI in healthcare.

5. Conclusion

Our comparative study demonstrates the potential of hybrid DL approaches, particularly the CNN-LSTM model, for accurate identification of chronic diseases. The high performance across multiple diseases and the ability to integrate diverse data types suggest that these models could be valuable tools for population health management and personalized preventive care.

The interpretability techniques employed provide insights into the models' decision-making processes, potentially facilitating their integration into clinical workflows. However, the observed performance variations across subgroups highlight the need for careful validation and potential adaptation before widespread clinical implementation. As healthcare continues to move towards precision medicine, hybrid deep learning models offer a promising approach to leveraging the wealth of available medical data for improved patient outcomes. Future research should focus on prospective validation, integration with existing clinical decision support systems, and assessment of the long-term impact on patient care and healthcare costs. External validation on diverse, multi-institutional datasets and Integration of genomic data are used to enhance predictive accuracy. Prospective studies to assess the impact of model-guided interventions on patient outcomes. Development of adaptive models can be fine-tuned to meet specific clinical settings.

References

[1] J. I. Z. Chen and P. Hengjinda, "Early prediction of coronary artery disease (CAD) by machine learning method—A comparative study," *J. Artif. Intell.*, vol. 3, no. 1, pp. 17–33, 2021.

[2] Y. Tong, A. I. Messinger, and G. Luo, "Testing the generalizability of an automated method for explaining machine learning predictions on asthma patients' asthma hospital visits to an academic healthcare system," *IEEE Access*, vol. 8, pp. 195971–195979, 2020.

[3] C. Wu, R. Gao, and Y. Zhang, "mHMDA: Human microbe-disease association prediction by matrix completion and multi-source information," *IEEE Access*, vol. 7, pp. 106687–106693, 2019.

[4] A. Kumar and H. Anjomshoa, "A two-stage model to predict surgical patients' lengths of stay from an electronic patient database," *IEEE J. Biomed. Health Inform*, vol. 23, no. 2, pp. 848–856, 2019.

[5] T. Wang, Y. Tian, and R. G. Qiu, "Long short-term memory recurrent neural networks for multiple diseases risk prediction by leveraging longitudinal medical records," *IEEE J. Biomed. Health Inform*, vol. 24, pp. 2337–2346, 2020.

[6] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, "HDPM: An effective heart disease prediction model for a clinical decision support system," *IEEE Access*, vol. 8, pp. 133034–133050, 2020.

[7] Y. Dong, Y. Sun, C. Qin, and W. Zhu, "EPMDA: Edge perturbation-based method for miRNA-disease association prediction," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 17, no. 6, pp. 2170–2175, 2020.

[8] Y. Zhao et al., "Prediction of Alzheimer's disease progression with multi-information generative adversarial network," *IEEE J. Biomed. Health Inform*, vol. 25, pp. 711–719, 2020.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025

[9] J. Luo and Y. Long, "NTSHMDA: Prediction of human microbe-disease association based on random walk by integrating network topological similarity," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 17, pp. 1341–1352, 2020.

[10] J. Xie and Q. Wang, "Benchmarking machine learning algorithms on blood glucose prediction for type I diabetes in comparison with classical time-series models," *IEEE Trans. Biomed. Eng.*, vol. 67, pp. 3101–3124, 2020.

[11] F. Ma et al., "Length-of-stay prediction for pediatric patients with respiratory diseases using decision tree methods," *IEEE J. Biomed. Health Inform*, vol. 24, no. 9, pp. 2651–2662, 2020.

[12] P. Wu et al., "An effective machine learning approach for identifying non-severe and severe coronavirus disease 2019 patients in a rural Chinese population: The Wenzhou retrospective study," *IEEE Access*, vol. 9, pp. 45486–45503, 2021.

[13] P. Chittora et al., "Prediction of chronic kidney disease—A machine learning perspective," *IEEE Access*, vol. 9, pp. 17312–17334, 2021.

[14] D. Gaglione et al., "Adaptive Bayesian learning and forecasting of epidemic evolution—Data analysis of the COVID-19 outbreak," *IEEE Access*, vol. 8, pp. 175244–175264, 2020.

[15] C. Wang et al., "Pulmonary image classification based on Inception-V3 transfer learning model," *IEEE Access*, vol. 7, pp. 146533–146541, 2019.

[16] D. V. K. Daliya, T. K. Ramesh, and S.-B. Ko, "An optimized multivariable regression model for predictive analysis of diabetic disease progression," *IEEE Access*, vol. 9, pp. 99768–99780, 2021.

[17] G. M. I. Sami Bourouis, M. H. Rashid, T. Tazin, and M. M. Khan, "Comparative analysis for prediction of kidney disease using intelligent machine learning methods," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–10, 2021.

[18] R. G. Nadakinamani et al., "Clinical data analysis for prediction of cardiovascular disease using machine learning techniques," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–13, 2022.

[19] M. S. Pathan, A. Nag, and S. Dev, "Analyzing the impact of feature selection on the accuracy of heart disease prediction," *Healthcare Anal.*, vol. 2, no. 100060, pp. 1–16, 2022.

[20] B. Nistal-Nuño, "Developing machine learning models for prediction of mortality in the medical intensive care unit," *Comput. Methods Programs Biomed*, vol. 216, no. 106663, pp. 1–18, 2022.

[21] B. Liao, X. Jia, and R. Sun, "DHDIP: An interpretable model for hypertension and hyperlipidemia prediction based on EMR data," *Comput. Methods Programs Biomed*, vol. 226, no. 107088, pp. 1–18, 2022.

[22] S. Huda et al., "A hybrid feature selection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis," *IEEE Access*, vol. 4, pp. 9145–9154, 2016.

[23] N. Sedaghat, M. Fathy, M. H. Modarressi, and A. Shojaie, "Combining supervised and unsupervised learning for improved miRNA target prediction," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 15, pp. 1594–1604, 2018.

[24] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.

[25] R. H. and T. Asha, "A linear model based on principal component analysis for disease prediction," *IEEE Access*, vol. 7, pp. 105314–105318, 2019.

[26] P. Sherubha, P. Amudhavalli, and S. P. Sasirekha, "Clone attack detection using random forest and multi-objective cuckoo search classification," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, 2019, pp. 450–454.

[27] S. Dinesh et al., "Investigations on brain tumor classification using hybrid machine learning algorithms," *J. Healthcare Eng.*, vol. 2022, pp. 1–10, 2022.

[28] P. Sherubha, "Graph-based event measurement for analyzing distributed anomalies in sensor networks," *Sådhanå (Springer)*, vol. 45, p. 212, 2020.

[29] M. H. Alnamoly, A. A. Hady, S. M. A. El-Kader, and I. M. El-Henawy, "A novel approach for enhancing image classification using deep learning techniques," *Int. J. Comput. Inf.*, vol. 3, no. 1, pp. 1–10, 2024.

[30] A. N. H. Zaied, M. Ismail, and S. El-Sayed, "An integrated knowledge management capabilities framework for assessing organizational performance," *Int. J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, pp. 1–10, 2012.

DOI: https://doi.org/10.54216/JCIM.160109
Received: November 20, 2024 Revised: January 28, 2025 Accepted: February 24, 2025