International Journal of Basic and Applied Sciences, 14 (SI-1) (2025) 122-130

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/mr62y398
Research paper

Revolutionizing Healthcare Analytics with A Robust Model for Secure Data Management and Superior Disease Prediction

S. Senthamarai [1]*, Dr. R. Mala [2], Dr. V. Palanisamy [3]

[1] Research Scholar, Department of Computer Applications, Alagappa University, Karaikudi, Tamilnadu, India
[2] Asst. Prof & Head, Department of Computer Science, Government Arts and Science College for Women, Paramakudi, Tamilnadu, India
[3] Senior Professor & Head, Department of Computer Applications, Alagappa University, Karaikudi, Tamilnadu. India.
**Corresponding author E-mail: sakthisiva2125@gmail.com*

Received: May 2, 2025, Accepted: May 22, 2025, Published: July 7 2025

Abstract

In healthcare, ensuring secure patient data management and leveraging predictive analysis are pivotal for enhancing medical diagnostics and treatment. The exponential growth in healthcare data, while fostering innovative solutions, raises concerns about data security and effective disease prediction. Traditional security approaches often fall short against sophisticated cyber threats, risking patient privacy. This research addresses these challenges comprehensively, proposing a model Particle Swarm-optimized Accentuate Attentive Layer Convo Recurrence Network (PS-AAL-CRNN) to safeguard patient data and advance disease prediction through sophisticated techniques. The research utilizes a healthcare dataset as the foundation for analysis and prediction. Preprocessing, this involves label encoding for categorical variables and robust scaling to mitigate the impact of outliers. Introduce data security by using hybrid encryption scheme, employing Blowfish-Elliptic Curve Diffie- Hellman (ECDH) for secure key exchange and Hash-Based Message Authentication Code (HMAC) for data integrity verification. Using Isomaptechnique to extracting essential features through nonlinear dimensionality reduction. For feature selection employ the Greedy Forward Feature Selection (GFFS) to optimize disease prediction by selectively identifying and retaining highly relevant features.Classification is performed using a PS-AAL-CRNN, with an attentive layer emphasizing critical features for precise disease prediction.Our model achieved better accuracy of 97.09%, precision of 97.97%, recall of 94.17%, f1-score of 96.03%, R2 of 0.843, PRC of 0.9883 with existing methods in performance evaluation.

Keywords: *Healthcare data, Security; Robust Scale; Blowfish-ECDH with HMAC; Isomap; Greedy Forward Feature Selection; Particle Swarm-optimized Accentuate Attentive Layer Convo Recurrence Network (PS-AAL-CRNN).*

1. Introduction

The healthcare industry, over the last few decades, has undergone a transformative shift, technology advancement and data science. The integration of digital tools with data analytics has enabled greater efficiency in patient care, increased accuracy in diagnostics, and more effective management of health-related data [1]. However, significant challenges in the healthcare industry persist: in terms of security and privacy over data, and the reliability of the predictive models deployed in disease diagnosis [2]. All these issues can be addressed with help of strong systems that integrate secure data management and advanced techniques in disease prediction to improve health outcome. Healthcare data is perhaps one of the most sensitive and valuable assets, that contain a wide range of information, such as patient medical histories, results of clinical trials, health records, and even medical images [3]. There are both opportunities and challenges associated with the massive amount of data produced every day by clinics, hospitals, wearable technology, and other medical equipment. A data analytics has the potential to revolutionize the healthcare industry by providing invaluable insights for predictive modeling, but data security and privacy are crucial[4]. The recurrent nature of cyber attacks and unauthorized access of data, secure data management frameworks are of principal importance to ensure safety for the sensitive information of patients without neglecting legal regulations [5]. Beyond protecting health care data, predictive analytics in health care has become a very powerful tool for early disease detection, development of personalized treatment plans, and better patient outcomes. The machine learning (ML) and artificial intelligence (AI) algorithms have proven to be exceptionally capable in analyzing large datasets to detect patterns and trends that might be impossible or difficult to identify by human practitioners [6]. Such models are able to predict how the disease will progress, thereby giving recommendations for the optimal treatment and even potential risks to a person's health, which might appear before a serious disease. Nevertheless, such impressive potential in the application of these technologies has to be optimized well for effectiveness in clinical application [7]. The integration of a robust model combining secure data management with advanced disease prediction algorithms can revolutionize healthcare analytics [8].While a predictive model can provide tailored treatment regimens, anticipate potential health hazards, and even expedite clinical procedures, a secure data management system safeguards all patient data. By allowing healthcare providers to make more precise and rapid data-driven decisions, the dual strategy can improve the outcome of the patients [9]. Superior disease prediction combined with safe data management has enormous potential to revolutionize the healthcare industry. With the right infrastructure and technologies in place, healthcare organizations can harness the full potential of data analytics to improve patient care, streamline operations, and ultimately revolutionize the way healthcare is delivered worldwide [10]. This model is to protect sensitive patient data throughout the healthcare system while exploiting advanced ML techniques to provide more accurate and timely predictions.

Contributions of this study

- The rapid expansion of healthcare data positionsrequests data security and accurate disease prediction, even as it encourages creative solutions.
- To protect patient data and improve disease prediction using advanced approaches, a model Particle Swarm-optimized Accentuate Attentive Layer Convo Recurrence Network (PS-AAL-CRNN) is proposed.
- A PS-AAL-CRNN is used for classification, with an attentive layer that highlights important characteristics for accurate disease prediction and proposed model outperforms current performance evaluation techniques.

The paper is organized into several sections: Part 2 covers the literature review, Part 3 and 4 methodology and outcomes, and Part 5 provides the study's conclusion.

2. Related works

The system gathers data from IoT devices and applies predictive analytics to patient history-related electronic clinical data that is stored on the cloud [11]. New technologies that allow cellular networks are used to gather data from sensor nodes, send it to local databases, and store it in cloud storage systems. The experimental findings are capable for the suggested prediction model for an M-Health system [12]. The growing needs of information and communication technology (ICT) applications, such as healthcare, smart cities, education, require sustainable energy-efficient networking architectures [13]. The intricacy of human brain tissues collective with the imaging process of nuclear magnetic resonance (NMR) outcomes in a multiplicity of noise, weak boundaries, and aberration in brain MRI (Magnetic Resonance Imaging) images. According to the findings, brain imaging feature recognition and predictive diagnosis can have an experimental foundation [14]. To develop dependable, cost-sensitive classifiers that can be utilized to accurately forecast medical diagnoses, some familiar algorithms, such as decision trees (DT), random forests (RF), extreme gradient boosting (GB), and logistic regression (LR), had their objective functions changed. The experiment's results show that cost-sensitive methods perform better than traditional algorithms [15]. To lower the number of deaths from cardiovascular disorders creates a model that can accurately forecast these conditions. An approach to k-mode clustering that can increase classification accuracy is presented. XGBoost (XGB), multilayer perceptron's (MP), RF, and decision tree classifiers (DT) are among the models that are employed [16]. An ideal ML model combined with extensive medical data on cardiac conditions can accomplish. Outcome demonstrates that, when compared to other cutting-edge techniques, methodology is more successful at predicting cardiac disease [17]. The suggested study's objective is to use ML techniques to determine the essential features of heart disease prediction. The simulation's outcomes show that the Modified Symbiotic Organisms Search -Adaptive Neuro-Fuzzy Inference System (MSSO-ANFIS) prediction model outperforms the alternative methods in terms of accuracy [18].

3. Methodology

The suggested PS-AAL-CRNNtechniques secure patient data management and enhanced disease prediction. A hybrid encryption scheme combining blowfish-ECDH and HMAC ensures data security, while Isomap and GFFS optimize feature extraction and selection. The model achieves superior performance in healthcare diagnostics through precise classification and robust security measures. Figure 1 shows the proposed flow.

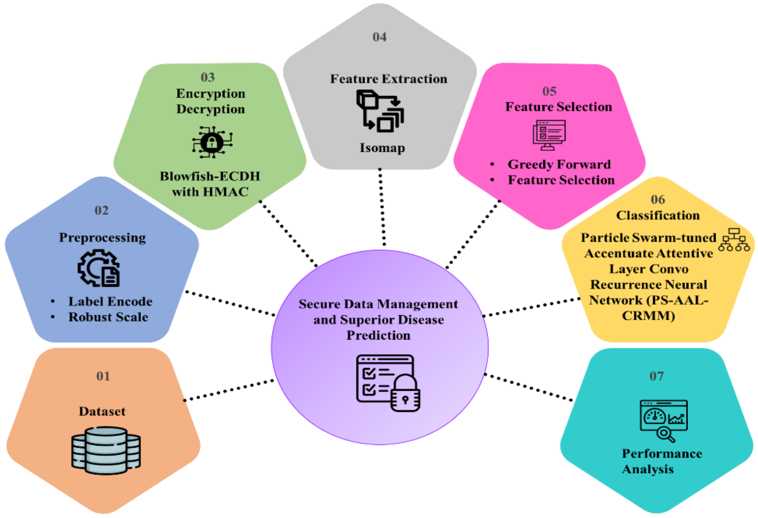

Fig. 1: Flow of Proposed.

3.1. Data collection

The collected data that includes four datasets from 1988: Cleveland, Hungary, Switzerland, and Long Beach V. A subset of 14 of the 76 traits, including the expected attribute, is mentioned in all published tests. The patient has heart disease, it is indicated in the "target" box. A number of 0 indicates no disease, while a value of 1 indicates disease. The data collected from Kaggle (https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset) website.

3.2. Data preprocessing

Preprocessing prepares data for analysis by addressing categorical and numerical differences. Label encoding converts categorical variables into numerical format, ensuring compatibility with Deep learning (DL) models. Robust scaling normalizes numerical data by focusing on the median and interquartile range, reducing the influence of outliers. Together, these steps enhance model accuracy and robustness.

3.2.1. Label encode

The sequence of categories under a certain feature attribute is used in label encoding; the category is substituted with sequential numbers ranging from 1 to n-1. This encoding technique is effective for health data management. But when applied to categorical variables in tasks improved anddisease prediction, high-based data that is not sized which could result in inaccurate model outputs and distort the model's interpretation of the categorical data, leading to lower accuracy.

3.2.2. Robust scale

A common method for performing robust scaling, first removing the median and then dividing by the interquartile range,equation (1) shows how this method reduces the impact of outliers by simplifying data management:

$$\mathrm{W_{scaled}} = \frac{\mathrm{W-median}}{\mathrm{R3(W)-R1(W)}} \tag{1}$$

It is particularly useful in Health data management applications, for example in the preprocessing of datasets towards better disease prediction where outliers can distort the outcome easily. It is a learned transformation; it learns appropriate values for median and inter-quartile range from training and applies them to data with which it performs transformations; and software implementations allow ranges to be changed. The approach reduces the impact of extreme values by focusing mainly on the core region of health care data. It mostly uses scaling method for variables that have outliers.

3.3. Encryption and decryption

This hybrid scheme employs Blowfish for the encryption of data and the ECDH scheme for secure key exchange and ensures strong security of the Health care data. Furthermore, the incorporation of the HMAC to verify integrity in data communication.

3.3.1. Blowfish-elliptic curve diffie-hellman (ECDH)

The symmetric key encryption is the Blowfish algorithm, which has a 64-bit block size and a key length that could be used to between 32 and 448 bits. The Blowfish algorithm's code is difficult to crack, the key size is larger. However, it is vulnerable to every attack, and save the low main class attack. When used in conjunction with Blowfish, ECDH can offer a safe way to exchange keys over an unprotected channel, guaranteeing that only authorized parties are capable of encryption and decryption.The Blowfish algorithm consists of two parts: data encryption and key expansion,the 448-bit main key is converted to 4168 bytes during the main expansion stage. The initialization is set to hexadecimal π digits for a P series of size 18 and four S boxes, each of which has a size of 256. Data encryption is done in 16 cycles. The Blowfish F function receives the result of XORing the leftmost 32 bits of plaintext with a 32-bit sub key in each round.

3.3.2. Hash-based message authentication code (HMAC)

The outcome of the MAC approach is named HMAC-X, in which X is the name of the hash function. SHA-2 and SHA-3 are examples of a variety of cryptographic hash functions which can be used to approximate HMAC. Hash output size, key size and quality, and hash function cryptographic strength influence the HMAC cryptographic strength. It ensures strong data integrity verification by using two passes of the corresponding hash estimation. HMAC display message integrity and authenticity, with any change to the message detectable, due to a secret key that is combined with cryptographic hashing in the context.

First, with the secret key, two keys are derived: the outer key and the inner key. At the first iteration, on the message, the inner hash and inner key are determined. The second pass, using an outer key and the result from the inner hash, gives the final HMAC code. Accordingly, the method is also more resistant to length-extension attacks. The iterative hash method and the input message into blocks based on a predetermined size. Using the HMAC by gatheringa knowledge of the secret key hashed and compared the HMAC in validation of the message's genuineness and integrity based on equations (2 and 3).

$$\mathrm{HMAC(Y,b)} = \mathrm{G}\left((\mathrm{Y}' \oplus \mathrm{O_{Outer}})||\mathrm{G}\left((\mathrm{Y}' \oplus \mathrm{O_{Outer}})||\mathrm{b}\right)\right) \tag{2}$$

$$\mathrm{Y}' = \begin{cases} \mathrm{G(Y), ForYvaluehigherthantheblocksize} \\ \mathrm{y, Otherwise} \end{cases} \tag{3}$$

3.4. Feature extraction using isomap

An approach to reducing dimensionality that is nonlinear, the IsoMap algorithm attempts to maintain the geodesic distances between data in a lower-dimensional environment. The approach improves the dimension reduction method of multi-dimensional scaling (MDS). First several steps are different between IsoMap and traditional MDS. This approach is utilized to data management as it makes processing complicated datasets easier. Moreover, IsoMap can help reduce the high-dimensionality of data when combined with better diseaseprediction models, making the precision and interpretability of predictive models in the medical domain more enhanced.

3.5. Feature selection using greedy forward

Feature selection is mainly applied in DL to extract only the most important qualities or optimal set of parameters to optimize the model's performance. This process is crucial to Health data management because it helps find and save the most important attributes for the optimization of the model. It employs a greedy search strategy, comparing every feature combination to the specified evaluation criterion. For example, the procedure helps to choose the most predictive factors for precise forecasts, which is helpful for betterdisease prediction.

3.6. Particle swarm-optimized accentuate attentive layer convo recurrence network (PS-AAL-CRNN)

ThePS-AAL-CRNN model combining PSO for fine-tuning, AAL for feature extraction and CRNN to process sequential information robustly. It would attain better accuracy rates in recognizing complex patterns in the Healthcare dataset.

3.6.1. Accentuate attentive layer (AAL)

The AAL builds upon the attention mechanism by accentuating the focus on critical parts of the input data. The AAL highlights key characteristics or time steps, whereas typical attention techniques let the model prioritize them. This enhances the model's capacity to identify intricate patterns and long-range connections in the data. Neural networks' attention mechanism may be expressed using equation (4):

$$\mathrm{Attention}(\mathrm{R},\mathrm{L},\mathrm{U}) = \mathrm{Softmax}\,\frac{\mathrm{RL}^{\mathrm{S}}}{\sqrt{\mathrm{c_l}}}\,\mathrm{U} \tag{4}$$

To accentuate the attention mechanism, introduce a scaling factor α, which emphasizes more relevant features in the sequence in equation (5):

$$\mathrm{Accentuated\ Attention}(\mathrm{R},\mathrm{L},\mathrm{U}) = \mathrm{softmax}\alpha.\frac{\mathrm{RL}^{\mathrm{S}}}{\sqrt{\mathrm{c_l}}} \tag{5}$$

Ris the query matrix, L is the key matrix,U is the value matrix, $\mathrm{c_l}$is the dimensionality of the key vector, α is the accentuation factor, which amplifies the attention given to key features.

3.6.2. Convo recurrence network (CRNN)

The first part introduces the current basic classification models, which include CNN-based and RNN-based classification model processes. These are significant in healthcare analytics, as they allow for an effective data organization process with an increased precision of the disease prediction. Then, introduce modified classification models, such as CRNN-based classification models to enhance predictive capabilities and streamline data processing for more effective results in healthcare.

Convolutional Neural Network (CNN): A feed forward neural network is generally characterized as a function equation (6).

$$\mathrm{Z} = \mathrm{E}(\mathrm{W},\theta) \tag{6}$$

$\mathrm{W} = \{\mathrm{w_1},\mathrm{w_2},\dots,\mathrm{w_m}\}$Is the collection of input, $\mathrm{Z} = \{\mathrm{z_1},\mathrm{z_2},\dots,\mathrm{z_m}\}$is the collection of output, E E is a model for a feed forward NN and θ is a collection of model parameters.A composition of layers can serve this purpose. Z is a collection of classes to function as a classification model. CNN is one type of feed forward NN. Extracting the local features of the healthcare data is the convolutional layer's objective equation (7).

$$\mathrm{Z_E} = \mathrm{Conv}(\mathrm{W},\theta_{\mathrm{CONV}}) \tag{7}$$

$\mathrm{Z_E}$is a collection of features that the convolutional layer has retrieved from W, and θ_{CONV} is a set of parameters in the convolutional layer equation (8).

$$\mathrm{Z_{DE}} = \mathrm{Pool}(\mathrm{Z_E},\theta_{\mathrm{Pool}}) \tag{8}$$

The pooling layer expresses the compressed features from $\mathrm{Z_E}$as $\mathrm{Z_{DE}}$. A group of parameters in the pooling layer is called θ_{Pool}. To integrate and categorize features, CNN additionally uses Softmax layers and fullyconnected layers as a classification model. Thus, $\mathrm{Z} = \mathrm{E}(\mathrm{W},\theta)$ of CNN can be expressed as equation (9).

$$\mathrm{Y} = \mathrm{Softmax}(\mathrm{FC}(\mathrm{Pool}(\mathrm{Conv}(\mathrm{X},\theta_{\mathrm{CONV}}),\theta_{\mathrm{Pool}}),\theta_{\mathrm{FC}})) \tag{9}$$

Recurrent Neural Network (RNN): RNN is mainly used in sequentially structured data, such as disease prediction and healthcare analytics. RNNs are also very important in Healthcare data management, because it processes temporal data efficiently to reveal patterns and trends that are helpful in advanced disease prediction models.

A type of RNN model called LSTM-RNN can actually solve the disappearance grade issue with recollection cell for holding long-lasting data. The LSTM-RNN model also possesses classification model with Softmax and fully-connected layers as depicted in equation (10).

$$\mathrm{Z} = \mathrm{softmax}(\mathrm{FC}(\mathrm{LSTM}(\mathrm{w},\theta_{\mathrm{LSTM}}),\theta_{\mathrm{FC}})) \tag{10}$$

The categorization in theresearch is to determine the greater disease prediction and data management.

Modified Classification Model CRNN: When it comes to utilizing features from spatial data, like images, CNN-based models perform well. CNN, cannot handle sequential data. Conversely, RNN-based models are modeling sequential data. The propose combine CNN and RNN to develop a new model, CRNN, based on their individual features. CNN and RNN can be combined to create a CRNN model by placing CNN in front of the RNN. One way to represent $\mathrm{Z} = \mathrm{E}(\mathrm{W},\theta)$ of CRNN is as follows equation (11).

$$\mathrm{Z} = \mathrm{softmax}(\mathrm{FC}(\mathrm{LSTM}(\mathrm{Pool}(\mathrm{Conv}(\mathrm{W},\theta_{\mathrm{Conv}}),\theta_{\mathrm{Pool}},)\theta_{\mathrm{LSTM}})\theta_{\mathrm{FC}})) \tag{11}$$

Another combination is to build an RCNN model by placing RNN in front of the CNN. One way to describe $\mathrm{Z} = \mathrm{E}(\mathrm{W},\theta)$of RCNN is as follows equation (12).

$$Z = \mathrm{softmax}(\mathrm{FC}(\mathrm{Pool}(\mathrm{Conv}(\mathrm{LSTM}(W, \theta_{\mathrm{LSTM}}), \theta_{\mathrm{Conv}}), \theta_{\mathrm{Pool}}), \theta_{\mathrm{FC}})) \tag{12}$$

Next, highlight and compress the text and temporal feature combination using the maximum pooling layer. The enhanced CRNN model operates on the premise that the problem is brief sequential equation (13).

$$Z = \mathrm{softmax}\big(\mathrm{FC}(\mathrm{Pool}(\mathrm{LSTM}(\mathrm{Conv}(W, \theta_{\mathrm{Conv}}), \theta_{\mathrm{LSTM}}), \theta_{\mathrm{Pool}}), \theta_{\mathrm{FC}})\big) \tag{13}$$

3.6.3. Particle swarm optimization (PSO)

The PSO method is used to optimize the parameters; they have a big impact on the expected outcomes flow. The PSO algorithm mimics the predatory behavior of birds through swarm intelligence optimization.By utilizing this algorithm, precise disease prediction models can be created, guaranteeing better decision-making and better health outcomes: wk=$(\mathrm{D}, \varepsilon, \gamma, \beta)$. The optimal solution of the vector's elements is the same as the global ideal position discovered throughout the optimization procedure. The approach is particularly valuable in healthcare analytics. By leveraging this algorithm, accurate disease prediction models can be developed, ensuring better decision-making and improved healthcare outcomes: wk= $(\mathrm{D}, \varepsilon, \gamma, \beta)$.

The present speed vector $u_k = \left(u_d, u_\varepsilon, u_\gamma, u_\beta\right)$decide their way and distance when movement. The updatelaw of particle is as follow equation (14 and 15).

$$u_k(s+1) = \theta u_k(s) + d_1 u_1\big(o_k - w_k(s)\big) + d_2 u_2\big(o_h - wk(s)\big) \tag{14}$$

$$w_k(s+1) = w_k(s) + u_k(s+1) \tag{15}$$

The local optimal value of the k-th particle is denoted by $o_k = \left(o_k^1, o_k^2, o_k^3, o_k^4\right)$, the global optimal value of the entire particle swarm by $\left(o_h^1, o_h^2, o_h^3, o_h^4\right)$, θ is the inertia weight, d_1 and d_2 are two positive constants that represent the local and global learning factors, and r_1 and r_2 are two random numbers that follow a uniform distribution in the interval [0,1].

The PSO-AAL-CRNN, integrating into the architectureenhanced attention and the PSO to enable focusing on important components through parameter optimization abilities. The architecture enhances DL feature selection and sequence modeling by using PSO to optimize the feature space, CNN for spatial relationships, RNN for temporal patterns, and an attention layer to amplify critical elements in the Health care data for disease prediction. Each component plays a role in making the model efficient, accurate, and able to handle complex, high-dimensional data.

4. Result

The model uses hybrid encryption, dimensionality reduction, and feature selection techniques to enhance patient data security and improve disease prediction by outperforming existing methods both in data protection and predictive accuracy. The development environment used Intel core i7 CPU running Windows 11, Python 3.8, 32 GB of RAM, and a JupyterNotebookfor interactive scripting and data processing.Comparison of our proposed method with the existing approaches such as secure K-medoids (SKM)[19], FNN-RLU [20] and AHT [21] in the metrics of accuracy, precision, recall, F1 score, and PRC. Table 1 depicts the parameters results.

Table 1: Outcomes of Metrics

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	R-Square	PRC
Secure k-medoids [19]	82.5	81.28	80.5		0.023	
FNN-RLU [20]	93.66			95.39		
AHT [21]	95.67	90.2		93.9		0.972
PS-AAL-CRNN [Proposed]	97.09	90.2	94.17	96.03	0.843	0.9883

4.1. Accuracy

A model's accuracy is a performance statistic that quantifies the percentage of accurate predictions it makes overall. In healthcare analytics, accuracy is essential for evaluating disease prediction models, as it reflects how well the system classifies patient conditions, disease risks, and health outcomes. By ensuring that these models are trained on relevant and high-quality information, effective data management enhances the performance and dependability of illness prediction systems. Figure 2 depicts the outcomes of accuracy.

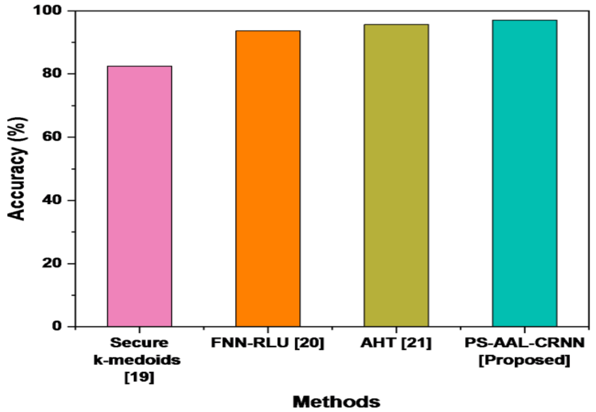

Fig. 2: Comparison of Accuracy.

The PS-AAL-CRNN methodaccuracy achieved 97.09% which is compared with the existing methods including SKM [19] achieved 82.50%, FNN-RLU [20] attained 93.66% and AHT [21] 95.67%.It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

4.2. Precision

Precision is a measure used in the evaluation of the accuracy of a classification model, particularly in healthcare analytics and disease prediction. It measures the proportion of true positive predictions made against all positive predictions. A greater accuracy reduces false positives, which increases the model's dependability in forecasting the positive class a crucial component in correctly categorizing diease. Figure 3 illustrates the results for precision.

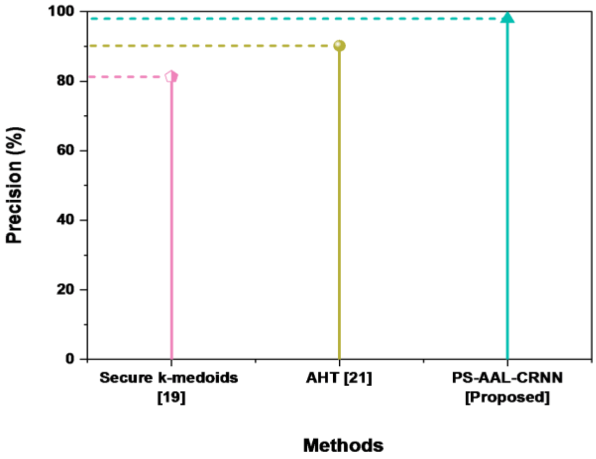

Fig. 3: Comparison of Precision.

The PS-AAL-CRNN achieved 97.97% precision which is compared with the existing methods including SKM [19] achieved 81.28% and AHT [21] 90.2%.It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

4.3. Recall

Recall is a measure of performance that is used for the evaluation of the adequacy of a model toward identifying the relevant instances or cases in healthcare analytics, disease prediction, and data management. It measures the number of correctly labeled positive instances divided by the total number of actual positive instances. For example, in medical diagnosis or fraud detection, such cases are of utmost importance; however, it could increase false positives. In any case where the minimization of missed detections is vital to making the correct predictions and decision-making, this metric becomes quite important. Figure 4 shows the recall.

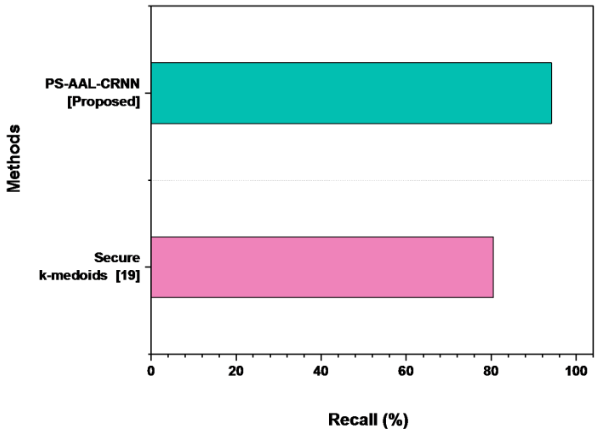

Fig. 4: Comparison of Recall.

The PS-AAL-CRNN methodrecall achieved 94.17% which is compared with the existing methods like SKM [19] achieved 80.50%.It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

4.4. F1- score

The F1 score is an important metric in healthcare analytics and disease prediction. It balances the precision and recall, hence giving an accurate assessment of a classification model's performance. Precision is the ratio of correct positive predictions made by the model, and recall is the ratio of actual positives correctly identified. In health care data management, F1 score aids in the optimization of the prediction of diseases especially where the cost of false positives and false negatives is crucially the same. F1 score varies between 0 and 1 with 1 representing perfect performance and 0 poor performances. Figure 5 describes the results of F1-score.

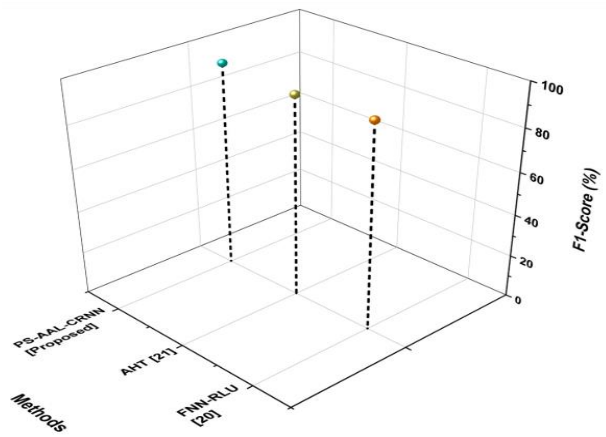

Fig. 5: Comparison of F1-Score.

The PS-AAL-CRNN methodaccuracy achieved 96.03% which is compared with the existing methods including FNN-RLU [20] attained 95.39% and AHT [21] 93.9%.It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

4.5. PRC

A binary classification model's performance is assessed using a graphical depiction called the Precision-Recall Curve (PRC). It displays the trade-off between recall (also known as sensitivity or true positive rate) and accuracy (also known as positive predictive value) over a range of categorization thresholds. In health care data management and disease prediction, the PRC is particularly valuable for evaluating models that identify rare conditions, such as cancer or heart disease, where achieving a balance between sensitivity (correctly identifying patients with the disease) and precision (minimizing false positives) is critical for effective diagnosis and treatment planning.Figure 6 and Figure 7 illustrates the results of PRC.

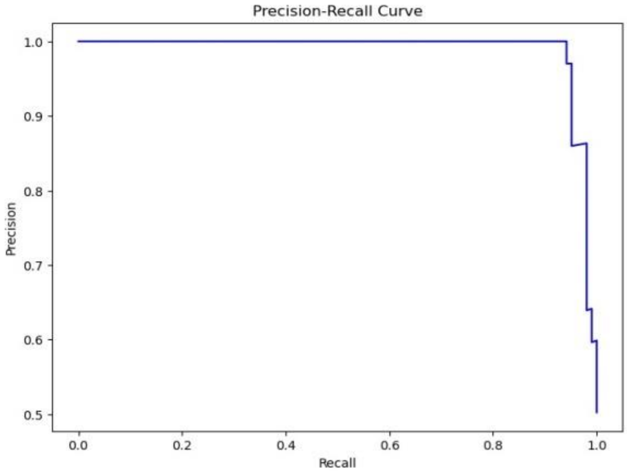

Fig. 6: Ouctome of PRC.

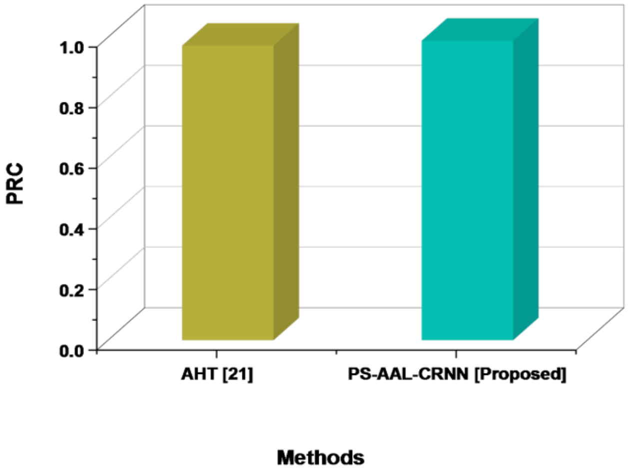

Fig. 7: Comparison of PRC.

The PS-AAL-CRNN 0.9883 PRC accuracy achieved which is compared with the existing method, AHT [21] 0.972. It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

4.6. R-Square

The term for R-Square (R^2) is a statistical measurement that indicates the proportion of variance in the dependent variable attributed to independent variables in the regression model. It defines goodness of fit of the regression model, used in healthcare analytics and the predictive potential of disease outcomes. Therefore, R^2 values ranging between close to 1 indicate good correlation and a better fit, while values of 0 indicate no explanatory power. Figure 8 in data management, the R^2 is useful for optimizing models that predict health trends or outcomes, guiding improvements in healthcare strategies and interventions.

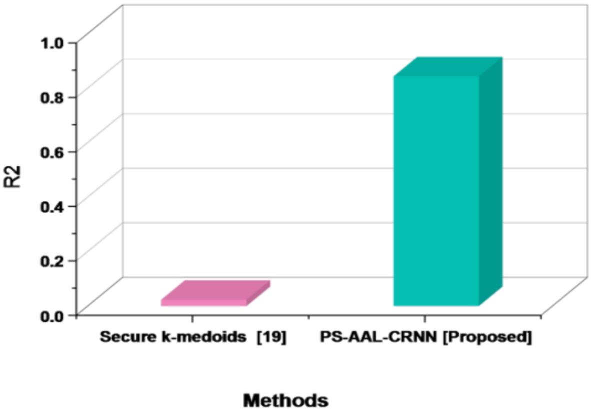

Fig. 8: Analysis of R^2.

The PS-AAL-CRNN methodR^2 achieved 0.843, which is compared with the existing methods, SKM [19] achieved0.023; It illustrates how the recommended methodology outperforms the existing approaches in terms of rates.

5. Discussion

The SKM method is sensitive to the initial choice of medoids, which can cause suboptimal clustering. It is also computationally expensive for large datasets due to its iterative nature. The FNN-RLU method is effective in some contexts but tends to suffer from over fitting, especially when the numbers of features are much larger than the number of samples. Besides this, it can demand the careful tuning of the network architecture to achieve optimal performance. Large-scale applications may find the AHT approach which is less effective due to its sluggish convergence and potential lack of scalability for high-dimensional data. For tasks such as image classification and time-series forecasting, the PS-AAL-CRNN approach is well-suited because it captures spatial features and temporal dependencies simultaneously. It's apyramid spatial attention technique focuses on the key areas, thus enhancing the accuracy. This can be achieved by using several model optimization techniques, such as quantization, pruning, or the use of lightweight designs that reduce computation while preserving performance.

6. Conclusion

A comprehensive approach to the dual challenge of protecting patient data and improving the predictive power of disease is proposed. Research propose the PS-AAL-CRNN by combining the state-of-the-art techniques for securing advanced data with the cutting-edge methods of machine learning for the confidentiality and accuracy of healthcare predictions.Hybrid encryption, dimensionality reduction, and feature selection are added to strengthen the ability of the model in providing secure handling of confidential information while enhancing more precise and reliable predictions about disease. In a comparative analysis,the suggested approach uses a number of criteria, including precision (90.2%), accuracy (97.09%), F1-score (96.03%) recall (94.17%), PRC(0.9883), and R^2 (0.843).

Limitations and future study

The model currently performs well with the data set, but further challenge in this model, scaling is more complex with healthcare datasets In the future, work could be done on diverse health care datasets to help better generalize and improve their accuracy. Data from many health care providers could allow seeing a greater understanding of disease patterns.

References

[1] Ismail, W.N., Hassan, M.M., Alsalamah, H.A. and Fortino, G., 2020. CNN-based health model for regular health factors analysis in internet-of-medical things environment. IEEE Access, 8, pp.52541-52549. https://doi.org/10.1109/ACCESS.2020.2980938.

[2] Ray, A. and Chaudhuri, A.K., 2021. Smart healthcare disease diagnosis and patient management: Innovation, improvement and skill development. Machine Learning with Applications, 3, p.100011. https://doi.org/10.1016/j.mlwa.2020.100011.

[3] Williams, R., Jenkins, D.A., Ashcroft, D.M., Brown, B., Campbell, S., Carr, M.J., Cheraghi-Sohi, S., Kapur, N., Thomas, O., Webb, R.T. and Peek, N., 2020. Diagnosis of physical and mental health conditions in primary care during the COVID-19 pandemic: a retrospective cohort study. The Lancet Public Health, 5(10), pp.e543-e550. https://doi.org/10.1016/S2468-2667(20)30201-2.

[4] Ahmed, I., Ahmad, M., Jeon, G. and Piccialli, F., 2021. A framework for pandemic prediction using big data analytics. Big Data Research, 25, p.100190. https://doi.org/10.1016/j.bdr.2021.100190.

[5] Guo, C. and Chen, J., 2023. Big data analytics in healthcare. In Knowledge technology and systems: Toward establishing knowledge systems science (pp. 27-70). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-99-1075-5_2.

[6] Khan, Z.F. and Alotaibi, S.R., 2020. Applications of artificial intelligence and big data analytics in m-health: A healthcare system perspective. Journal of healthcare engineering, 2020(1), p.8894694. https://doi.org/10.1155/2020/8894694.

[7] Peiffer-Smadja, N., Rawson, T.M., Ahmad, R., Buchard, A., Georgiou, P., Lescure, F.X., Birgand, G. and Holmes, A.H., 2020. Machine learning for clinical decision support in infectious diseases: a narrative review of current applications. Clinical Microbiology and Infection, 26(5), pp.584-595. https://doi.org/10.1016/j.cmi.2019.09.009.

[8] Alowais, S.A., Alghamdi, S.S., Alsuhebany, N., Alqahtani, T., Alshaya, A.I., Almohareb, S.N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H.A. and Al Yami, M.S., 2023. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC medical education, 23(1), p.689. https://doi.org/10.1186/s12909-023-04698-z.

[9] Ajegbile, M.D., Olaboye, J.A., Maha, C.C. and Tamunobarafiri, G., 2024. Integrating business analytics in healthcare: Enhancing patient outcomes through data-driven decision making. World J Biol Pharm Health Sci, 19, pp.243-50. https://doi.org/10.30574/wjbphs.2024.19.1.0436.

[10] Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, pp.208-222. https://doi.org/10.1016/j.inffus.2020.06.008.

[11] Nancy, A.A., Ravindran, D., Raj Vincent, P.D., Srinivasan, K. and Gutierrez Reina, D., 2022. Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. Electronics, 11(15), p.2292. https://doi.org/10.3390/electronics11152292.

[12] Qureshi, K.N., Din, S., Jeon, G. and Piccialli, F., 2020. An accurate and dynamic predictive model for a smart M-Health system using machine learning. Information Sciences, 538, pp.486-502. https://doi.org/10.1016/j.ins.2020.06.025.

[13] Bharathi, R., Abirami, T., Dhanasekaran, S., Gupta, D., Khanna, A., Elhoseny, M. and Shankar, K., 2020. Energy efficient clustering with disease diagnosis model for IoT based sustainable healthcare systems. Sustainable Computing: Informatics and Systems, 28, p.100453. https://doi.org/10.1016/j.suscom.2020.100453.

[14] Hu, M., Zhong, Y., Xie, S., Lv, H. and Lv, Z., 2021. Fuzzy system based medical image processing for brain disease prediction. Frontiers in Neuroscience, 15, p.714318. https://doi.org/10.3389/fnins.2021.714318.

[15] Mienye, I.D. and Sun, Y., 2021. Performance analysis of cost-sensitive learning methods with application to imbalanced medical data. Informatics in Medicine Unlocked, 25, p.100690. https://doi.org/10.1016/j.imu.2021.100690.

[16] Bhatt, C.M., Patel, P., Ghetia, T. and Mazzeo, P.L., 2023. Effective heart disease prediction using machine learning techniques. Algorithms, 16(2), p.88. https://doi.org/10.3390/a16020088.

[17] Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, pp.208-222. https://doi.org/10.1016/j.inffus.2020.06.008.

[18] Khan, M.A. and Algarni, F., 2020. A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS. IEEE access, 8, pp.122259-122269. https://doi.org/10.1109/ACCESS.2020.3006424.

[19] Akter, S., Reza, F. and Ahmed, M., 2022. Convergence of Blockchain, k-medoids and homomorphic encryption for privacy preserving biomedical data classification. Internet of Things and Cyber-Physical Systems, 2, pp.99-110. https://doi.org/10.1016/j.iotcps.2022.05.006.

[20] Vedaraj, M. and Ezhumalai, P., 2022. A Secure IoT-Cloud Based Healthcare System for Disease Classification Using Neural Network. Computer Systems Science & Engineering, 41(1). https://doi.org/10.32604/csse.2022.019976.

[21] Elbasi, E. and Zreikat, A.I., 2023. Heart Disease Classification for Early Diagnosis based on Adaptive Hoeffding Tree Algorithm in IoMT Data. Int. Arab J. Inf. Technol., 20(1), pp.38-48. https://doi.org/10.34028/iajit/20/1/5.

[22] Rimada, Y., Mrinh, K.L., &Chuonghan. (2024). Unveiling the printed monopole antenna: Versatile solutions for modern wireless communication. National Journal of Antennas and Propagation, 6(1), 1–5. https://doi.org/10.31838/NJAP/06.01.01.

[23] Uvarajan, K. P. (2024). Integration of artificial intelligence in electronics: Enhancing smart devices and systems. Progress in Electronics and Communication Engineering, 1(1), 7–12. https://doi.org/10.31838/ECE/01.01.02.

[24] Uvarajan, K. P. (2024). Advanced modulation schemes for enhancing data throughput in 5G RF communication networks. SCCTS Journal of Embedded Systems Design and Applications, 1(1), 7-12. https://doi.org/10.31838/ESA/01.01.02.