

REVIEW



An artificial flora algorithm-based recurrent neural network for multi-disease prediction in healthcare monitoring system

Anooj P. K

Department of Computing and Information Sciences, University of Technology and Applied Sciences-AI Musannah, Al Muladdah, Oman

ABSTRACT

Background: The growing patient population in today's digitally transformed healthcare environment is pushing the industry toward more advanced remote medical monitoring solutions. The capacity of these systems to accurately foresee health conditions is essential to their effectiveness. While widely used, conventional illness prediction models—which primarily depend on machine learning techniques—often have poor predictive accuracy.

Method: Our study presents a novel solution to this shortcoming: an improved Recurrent Neural Network (RNN) model that is further optimized by using the Artificial Flora (AF) technique. Through the careful use of the AF algorithm to optimize the weight parameter of the RNN, this novel technique greatly improves the model's performance in multi-disease prediction.

Findings: Our study includes a thorough assessment and contrast of the suggested RNN-AF model with other well-known models, such as the RNN and RNN- Particle Swarm Optimization (RNN-PSO). Key performance parameters including accuracy, sensitivity, and specificity provide a solid foundation for this comparison research.

Interpretation: Our tests' empirical results highlight the RNN-AF model's better performance and demonstrate how well it outperforms other models in terms of illness prediction accuracy.

KEYWORDS

Recurrent neural network;
Artificial flora algorithm;
Particle swarm optimization;
Kidney disease; Diabetes

ARTICLE HISTORY

Received 03 June 2024;
Revised 20 June 2024;
Accepted 27 June 2024

Introduction

Understanding the correlations between many elements, such as the patient's history, symptoms, and surroundings, is essential in the healthcare industry [1]. The healthcare frameworks are designed to fulfill the prerequisites of the increasing population worldwide. Individuals are impacted by life-threatening illnesses. Diabetes contributes to severe complications like heart attacks, kidney failure, and vision impairment. Because of the extensive effects of such diseases, healthcare systems worldwide utilize monitoring frameworks for numerous diseases and symptoms, by collecting data transmitted by patient monitoring devices integrated with Internet of Things (IoT) devices for remote monitoring. The deployment of knowledge and data mining strategies is essential to support these sophisticated systems. The administration, storage, integration, and interpretation of massive volumes of data must be handled with efficiency, and this requires the use of certain approaches [2].

IoT technology is a game-changer, leading to the creation of smart medical applications. This has revolutionized healthcare industry, making it a leading domain in adopting advanced technologies [3]. The ability to monitor health remotely using technologies like IoT, Edge-of-Things, and Cloud-of-Things, significantly reduces the need for hospital admissions [4]. Notably IoT, enables a host of applications in information technology and advanced, interconnected healthcare [5].

Recently, IoT has taken mobile healthcare (m-healthcare) to new heights, offering diverse services and dimensions online [6-9]. These applications are providing a new, accessible platform for millions seeking regular health advice, essentially promoting healthier living habits.

Wearable IoT devices are crucial for health tracking, measuring heart rate and body temperature, displaying data on mobile phones. This enables users to monitor their health effectively [10]. Patient data from these devices is analyzed with machine Learning (ML) to extract insights and predict health status.

ML techniques, like Hidden Markov Models, Neural Networks, and Support Vector Machines, along with Data Mining Techniques like Fuzzy Logic and Decision Trees (DT), are used in remote healthcare monitoring systems to predict diseases [11-15]. Additionally, the numerous ML algorithms used have improved disease prediction to trustable levels [16,17]. Yet, there exists scope for new techniques to further improve the accuracy.

The limitations in analyzing remote monitoring data using ML, is the complexity in understanding the unseen data, specifically by the complex Random Forests (RF) and DT. Training and optimizing the model becomes increasingly difficult with its complexity. Additionally, there exists a chance of overfitting, where the model fails to generalize adequately to new, unknown data [18]. These models have not delivered sufficient accuracy, owing to non-performing factors with the new data [19].

To enhance precision, we emphasize prediction models utilizing deep learning (DL). A novel approach called Artificial Flora Algorithm (AFA) is introduced. AFA is based on Recurrent Neural Network (RNN), utilizing which drawbacks of ML models in generalizing new data can be overcome by updating/optimizing the weight parameters.

*Correspondence: Dr. Anooj P. K, Department of Computing and Information Sciences, University of Technology and Applied Sciences-AI Musannah, Musannah, Oman, e-mail: anooj.pk@utas.edu.om

© 2024 The Author(s). Published by Reseapro Journals. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

AFA may be utilized in healthcare monitoring systems that employ DL and neural network approaches to overcome the shortcomings of multi-disease prediction models. AFA is made to address difficult, discrete, nonlinear optimization issues, drawing inspiration from migration and reproduction strategies of flora. With DL models like RF or DT, the AFA might be useful for streamlining feature extraction and data normalization procedures, decreasing overfitting, and optimizing model parameters including weights.

The contribution of the paper is described as follows

- For healthcare prediction, a novel approach, the RNN model based on AFA is employed by optimizing the weight parameters of RNN. Indeed, AFA is harnessed for the optimization of the weight parameters within the RNN.
- In this study, datasets of kidney disease, diabetes, and heart disease are employed for healthcare prediction.
- The suggested approach is implemented using MATLAB.
- The effectiveness of the proposed prediction model is evaluated based on accuracy, sensitivity, and specificity.

The subsequent sections of this article are structured as follows: Section 2 delves into prior literature that primarily focuses on disease prediction research within healthcare monitoring systems. Section 3 proposes the AF algorithm-based RNN for Multi-Disease Prediction in Healthcare Monitoring System. In section 4, the proposed model's results are discussed. Finally, the conclusion of this paper is described.

Literature survey

The emergence of artificial intelligence (AI) and ML has brought a transformative shift in disease prediction in health surveillance systems. This review adopts a thematic approach to survey relevant literature, providing a critical analysis of different methodologies and linking them to specific research objectives.

Disease prediction using neural networks

Duneja et al. (2019) utilized LSTM networks to predict diseases, leveraging their ability to capture temporal processes [20]. While effective, this approach is computationally demanding and reliant on extensive training data, potentially limiting scalability [20]. In ophthalmology, Srinivas et al. devised a method for detecting age-related macular degeneration (AMD) using AMD-ResNet, a convolutional neural network, and the Internet of Medical Things [21]. Their study introduces a teleophthalmology framework employing a head-mounted camera to send retinal images to cloud storage. However, concerns over privacy in cloud-based systems and the necessity of significant computing resources for high-resolution image analysis may restrict its application.

A significant medical AI breakthrough involves an ensemble model merging ML and DL for illness prediction. This model showcases effective integration and optimization through validation loss, performing well on various metrics like recall, accuracy, precision, and F1-score. Applying these findings in practical clinical settings is difficult due to the diverse illnesses addressed, including acute viral hepatitis B and malaria [21]. This emphasizes the importance of flexibility and customization in medical settings. Future research opportunities are presented, especially to enhance the model's interpretability and expand its therapeutic usefulness.

Fuzzy systems and support vector machines

Vidhya and Shanmugalakshmi have introduced Modified Adaptive Neuro-Fuzzy Inference System (M-ANFIS) which deals with big data preprocessing in healthcare [22]. By extracting features and clustering closed frequent item sets (CFI), the methodology addresses the complexity of big data. This system is promising, but generalizing it to different datasets remains uncertain.

Harimoorthy and Thangavelu used improved SVM radial bias kernel technique for predicting chronic diseases by reducing the feature set [23]. This method is consistent with our goals of emphasizing computational efficiency and model interpretability. However, questions regarding the model's robustness and possible overfitting require further validation.

Ensemble and DL models

Ali et al. demonstrated the effectiveness of ensemble DL for heart disease prediction, such as stroke, heart attack, and hypertension, using a novel conditional probability strategy for feature weighting [24]. Managing High-Dimensional Datasets, with the complexity and volume of heart disease-related datasets, frequently poses challenges for traditional systems. The method efficiently handles high-dimensional datasets [24]. Selection of Features and Weighting using conventional techniques are usually employed by current systems to pick features and give them a general weight. They often lack optimal performance for diagnosing cardiac disease. The performance is improved using information gain for more effective feature selection and a conditional probability technique for accurate feature weighing [24]. Specificity measures in feature selection of heart disease-related data can improve the performance of predictive models, consistent with developing accurate predictive systems. However, the complexity of such models can make clinical implementation difficult in terms of interpretability and computational complexity.

Min Chen et al. demonstrated a multimodal disease risk prediction algorithm based on Convolutional Neural Networks (CNN), which demonstrated higher accuracy and faster convergence than unimodal algorithms using hospital data, highlighting the potential of utilizing structured and unstructured data for disease prediction [25]. Despite its success, its adaptability to different healthcare settings requires further investigation.

For diabetes prediction, Kumar et al. used a novel systematic approach to generate relevant medical data by leveraging the UCI repository dataset and medical sensors [7]. This strategy aimed to improve prediction accuracy of severe diabetes and integrate traditional dataset analysis with real-time physiological monitoring.

Materials and Method

Overview

A prediction model for multiple diseases, including kidney disease, diabetes, and heart disease, is introduced using the AFA-based RNN model, proposed as a prediction model for healthcare monitoring. Multi-disease datasets are input into the prediction model and analyzed. To optimize performance, the weight parameters of the RNN must be enhanced. Hence, these parameters are carefully selected by utilizing AFA. Initially, the model is trained by employing the AF-based RNN model. Subsequently, the system undergoes testing to assess its

performance. During the testing phase, it predicts normal or abnormal conditions. Figure 1 illustrates the overview of the proposed model.

Prediction model

RNN is used for classification, prognosis, and recognition. The RNN classifier processes input datasets to classify data as semi-spiral or non-spiral. The system involves two key phases: training and testing. It has three layers: input, hidden, and output layer. The hidden layer includes hidden and contextual layers. Notably, the one-step delays within the feedback path render the methodology akin to a feed-forward network. For the hidden layer, the results are utilized for feedback signals. To enhance the RNN, the weight parameters are upgraded using AFA. Figure 2, depicts the basic structure of RNN.

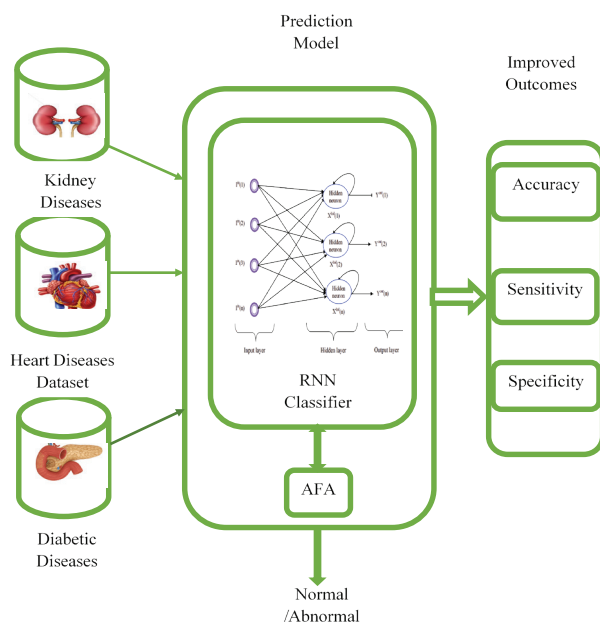


Figure 1. Overview of proposed prediction model.

The training process for the RNN involves using training samples in pairs (a_1, b_1) , (a_2, b_2) , ..., (a_n, b_n) , where a_i belongs to R^n and b_i belongs to R^n for values of i ranging from 1 to n . Here, a_i represents the input vector, and b_i represents the target output.

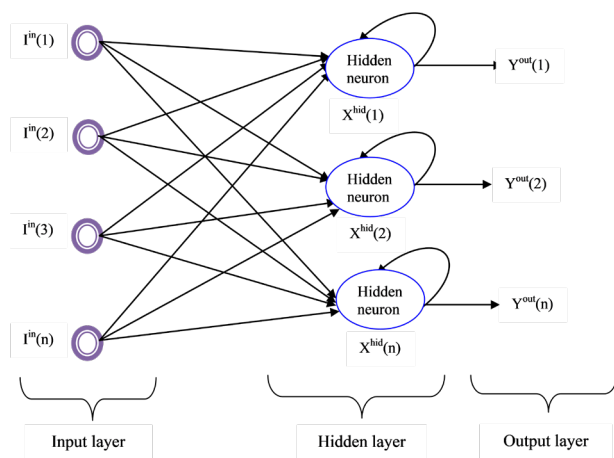


Figure 2. The fundamental structure of the recurrent neural network.

The steps are given below:

Step (1): First, the extracted feature dataset is provided as input to the RNN classifier. Step (2): Conventionally, the RNN can be described using the following equations (1) and (2).

$$x_i(t) = \sum_j y_j(t) w_{ij}(t) \quad (1)$$

$$y_i(t) = f_i(x_i(t)) \quad (2)$$

In various contexts, it is crucial to concurrently optimize both the neuron's initial method and the associated weights. The initial function relies on inputs from both the network layer and the context layer.

Step (3): The vector result determines the activation function of the hidden nodes, and this function is defined using the sigmoid as presented in equation (3).

$$f_i = \frac{1}{1 + e^{(-x_i)}} \quad (3)$$

For $i = 1$ and 2 , the RNN output is denoted as Y^{act} and is computed as $Y^{act} = W_{2i} f_i$, using a single output weight matrix.

Step (4): The release function of each neuron is computed using equations (4) and (5) ahead of the dissemination of the practice.

$$y_i(t) = f_i(x_i(t), C_i(t)) \quad (4)$$

$$x_i(t) = \sum_{j \in H} y_j(t) w_{ij} + \sum_{j \in I} x_j(t) w_{ij} + \sum_{j \in C} y_j(t - \tau_{ij}) w_{ij} \quad (5)$$

In this context, f , H , I , and C represent neuron functions with hidden layer, input, and final phase values. x_j is an input, τ_{iji} is an integer in the displacement series, and t indicates backpropagation error in neural computation.

Step (5): The backpropagation error is calculated using equation (6),

$$E_m = Y^{tar} - Y^{act} \quad (6)$$

To reduce an error significance, the weight factors of the RNN are modified using the AF method.

Weight updating using AF algorithm

AF algorithm

Mechanical spreading is autochory, while biological growth, anemochory, and hydrochory fall under allochory. Autochory enables plants to move to suitable habitats independently, whereas allochory allows plants to disperse to remote and unknown locations. These strategies expand the study of flora and reduce extinction risks. Flora distribution can change due to factors like climate, harsh environments, and competition. Plants, lacking mobility, rely on seed dispersal and growth to thrive. Evolution and reproduction help them find ideal habitats. Cheng et al. introduced the AF method, a swarm intelligence-based optimization technique, to enhance RNN weight values in response to such situations [26].

The AF method comprises four key components: the parent plant, the offspring plant, the location, and the propagation distance. This algorithm exhibits three primary behavioral characteristics: evolution, spreading, and selection. Evolution behavior reflects the plant's adaptation. Spreading behavior involves seed movement, through autochory or allochory. Selective behavior determines plant survival or extinction (Figure 3).

Optimization with the AF algorithm

The optimization process is as follows:

Initializing

The AF method produced the initial flora randomly using N options (or weight parameters for the RNN), resulting in N plants. The original plant positions are represented by a matrix called $P_{i,j}$, here, j represents the quantity of plants in the flora, and i denotes the dimension's size.

$$P_{i,j} = \text{rand}(0,1) \times d \times 2 - d \quad (7)$$

Where "rand(0, 1)" represents a set of randomly distributed integers within the range (0, 1), and "d" is the magnitude of the boundary area.

Evolution behavior

The parent plant disperses seeds in all directions within a circular area defined by its propagation distance, which evolved based on the parent and grandparent plant's propagation distances.

$$d_j = d_{1j} \text{rand}(0,1) \times c_1 + d_{2j} \times \text{rand}(0,1) \times c_2 \quad (8)$$

The learning coefficients are c_1 and c_2 , the parent plant's propagation distance is denoted as d_{2j} , and the grandparent plant's propagation distance is represented as d_{1j} . The independent evenly dispersed number in (0,1) is denoted as rand (0,1).

The propagation distance of the new grandparents is

$$d'_{1j} = d_{2j} \quad (9)$$

The spread distance of a fresh parent is the typical variation across the original and offspring plant sites.

$$d'_{2j} = \sqrt{\frac{\sum_{i=1}^N (P_{i,j} - P'_{i,j})^2}{N}} \quad (10)$$

Spreading behavior

The propagation function computes the position of the offspring plant using the following procedure:

$$P'_{i,j \times m} = D_{i,j \times m} + P_{i,j} \quad (11)$$

$P_{i,j}$ represents the present position of the parent plant, while $P'_{i,j \times m}$ indicates the location of the offspring plant. Here, m signifies the number of seeds that a single plant can disperse, and $D_{i,j \times m}$ represents a randomly generated integer following a Gaussian distribution with a mean of 0 and a variance of d_j .

Fitness

Once the solution process commences, the fitness score for each solution is computed to determine the ideal solution or weight parameter. The greatest predicted accuracy is regarded as the best fitness metric. The function of fitness is defined by equation (12)

$$\text{Fitness} = \max(\text{Accuracy}) \quad (12)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

Here, true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are used.

Select behavior

The likelihood of the offspring plants surviving is determined by the survival probability, which is defined as follows.

$$p = \sqrt{\frac{F(P'_{i,j \times m})}{F_{\max}}} \times Q_x^{(j \times m - 1)} \quad (14)$$

Considering $Q_x^{(j \times m - 1)}$ as Q_x raised to the power of $(j \times m - 1)$, where Q_x represents the selection probability and must fall

within the range of 0 to 1. It's worth noting that offspring plants situated farther from the original plant may exhibit reduced fitness levels. Q_x reflects the algorithm's exploratory capacity, with a need for higher values in scenarios involving problems with straightforward local optimal solutions. $F(P'_{i,j \times m})$ represents the fitness level of the j th solution, while $F_{\max}$ denotes the highest fitness level within the flora for the current generation.

Termination

The roulette wheel selection, also known as percentage selection, determines offspring plant survival based on probabilities, considering potential scores for each option. Decision-making is influenced by acknowledged likelihood and not solely by score value.

Testing stage

After training, we test the models to evaluate their prediction performance using 20% of the dataset. This helps determine if patients' conditions are categorized as normal or pathological.

Algorithm: Weight parameters of RNN selection using AFA

- **Step 1:** Set the number of initial plants (N) (or weight parameters), the maximum number of repetitions, and the population, as well as any other associated factors.
- **Step 2:** Determine the propagation distance using equations (8), (9), and (10).
- **Step 3:** The offspring plant's location is calculated using the equation (11).
- **Step 4:** Use equation (12) to get the fitness value.
- **Step 5:** Use the equation (14) to calculate the survival probability. The roulette wheel's selection mechanism determines whether or not the child survives.
- **Step 6:** If any plants survive, select at random N plants to become original plants; if none survive, produce new original plants.
- **Step 7:** Note the most effective weight parameter.
- **Step 8:** Repeat the procedure until the solution has the best fitness value, weight parameter, or number of iterations.
- **Step 9:** The process is halted after the optimum solution or weight parameter has been obtained.

The algorithm translation to flow chart is described in Figure 3.

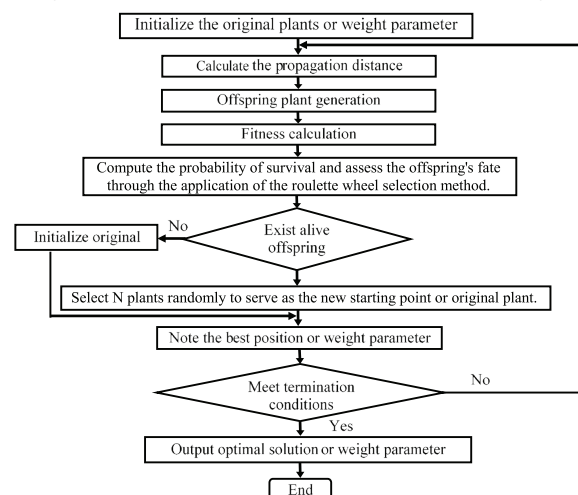


Figure 3. The flowchart of the AF algorithm.

Result and Discussion

We analyzed the performance of our proposed prediction model using various datasets, including those related to kidney disease, diabetes, and heart disease. Implementation of our system takes place on a computing platform equipped with an Intel(R) Core (TM) i5 processor, 4GB of RAM, and MATLAB 7.10.0 (R2015a). We assess the model's effectiveness by measuring accuracy, sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), False Positive Rate (FPR), False Negative Rate (FNR), and False Discovery Rate (FDR). Furthermore, we conduct comparative evaluations between our proposed prediction model and other existing models, including RNN and RNN- Particle Swarm Optimization (RNN-PSO).

Analysis of data on various diseases

Comprehensive description of the kidney disease dataset

This data is obtained from the National Institute of Diabetes, Digestive and Kidney Diseases (NIDDKD), India, collected over two months, including 25 characteristics (ID, age, blood pressure, number of red blood cells, white blood cell count, etc.). The dataset's objective is to predict Chronic kidney disease (CDK). A total of 400 rows was obtained.

Comprehensive description of the diabetes dataset

This data was obtained from the NIDDKD. The dataset's objective is to predict diabetes using the exact diagnostic

metrics provided in it. Numerous restrictions were posed in selecting cases from a large database. Specifically, the patients are Pima Indian girls over the age of 21. The datasets include medical predictor factors and one goal variable, Outcome. The predictive factors include the patient's age, body mass index, insulin level, and previous pregnancies.

Comprehensive description of the Heart Disease Dataset

The datasets for Cleveland, Hungary, and Switzerland are sourced from the UCI repository. These datasets include 76 characteristics. However, just 14 characteristics (age, sex, cp, trestops, chol, FBS, restecg, thalach, exang, oldpeak, slope, ca, thal, and num) are utilized in investigations.

Comparative analysis

To demonstrate the suggested model's usefulness, it is studied and compared to RNN and RNN-PSO models regarding accuracy, sensitivity, specificity, PPV, NPV, FPR, FNR, and FDR.

Analysis of dataset related to kidney disease

The proposed model excels in predicting renal disease, surpassing existing models in key performance metrics. Table 1 and Figure 4 detail the model's superior accuracy, sensitivity, and specificity. It shows near-perfect accuracy (0.995), high sensitivity (0.98), and unmatched specificity (1), with strong PPV and NPV values. Additionally, it maintains low FPR, FNR, and FDR rates, confirming its reliability over traditional models.

Table 1. Performance analysis of kidney disease dataset.

	Sensitivity	Specificity	Accuracy	PPV	NPV	FPR	FNR	FDR
RNN	0.8	1	0.95	1	0.9375	0	0.2	0
RNN-PSO	1	0.846667	0.885	0.684932	1	0.153333	0	0.315068
RNN-AFO	0.98	1	0.995	1	0.993377	0	0.02	0

Table 1 compares three ML models for predicting renal disease. RNN-AFO performs better, obtaining near-perfect sensitivity, accuracy, and specificity. The RNN model excels in specificity but lacks sensitivity. Conversely, RNN-PSO is highly sensitive but sacrifices specificity and accuracy. Both models exhibit low false discovery rates.

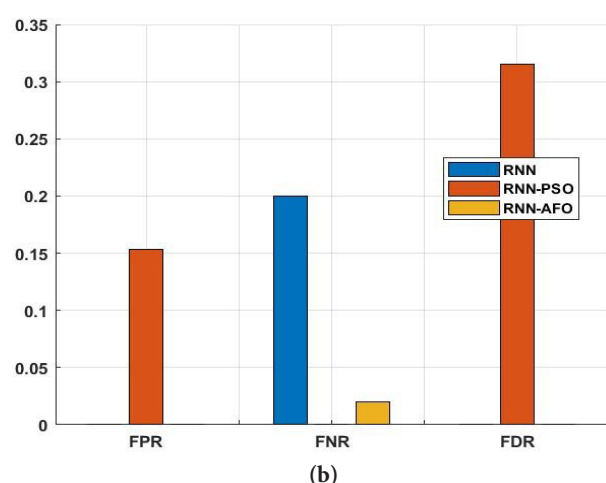
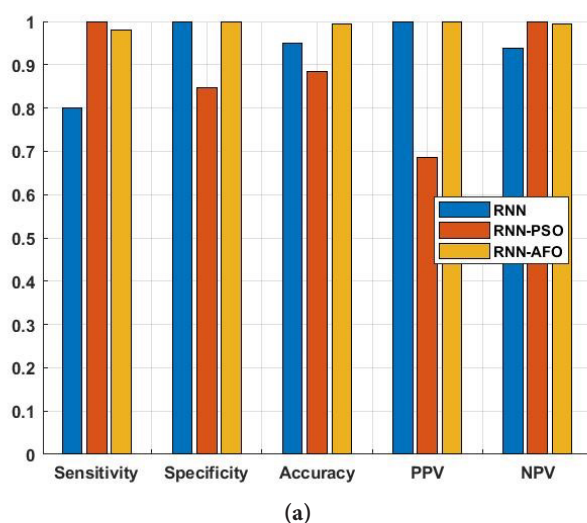


Figure 4 (a and b). The comparative analysis of different prediction models for kidney disease dataset.

Analysis of dataset related to diabetes

The model for predicting diabetes is more accurate than existing models, with significant improvements in accuracy (0.981771), sensitivity (0.988506), and specificity (0.96748). Table 2 and Figure 5 illustrate its superior performance across multiple metrics, and overall effectiveness in diabetes prediction.

Table 2. Performance analysis of kidney disease dataset.

	Sensitivity	Specificity	Accuracy	PPV	NPV	FPR	FNR	FDR
RNN	0.835249	0.463415	0.716146	0.767606	0.57	0.536585	0.164751	0.232394
RNN-PSO	0.846743	0.926829	0.872396	0.96087	0.74026	0.073171	0.153257	0.03913
RNN-AFO	0.988506	0.96748	0.981771	0.984733	0.97541	0.03252	0.011494	0.015267

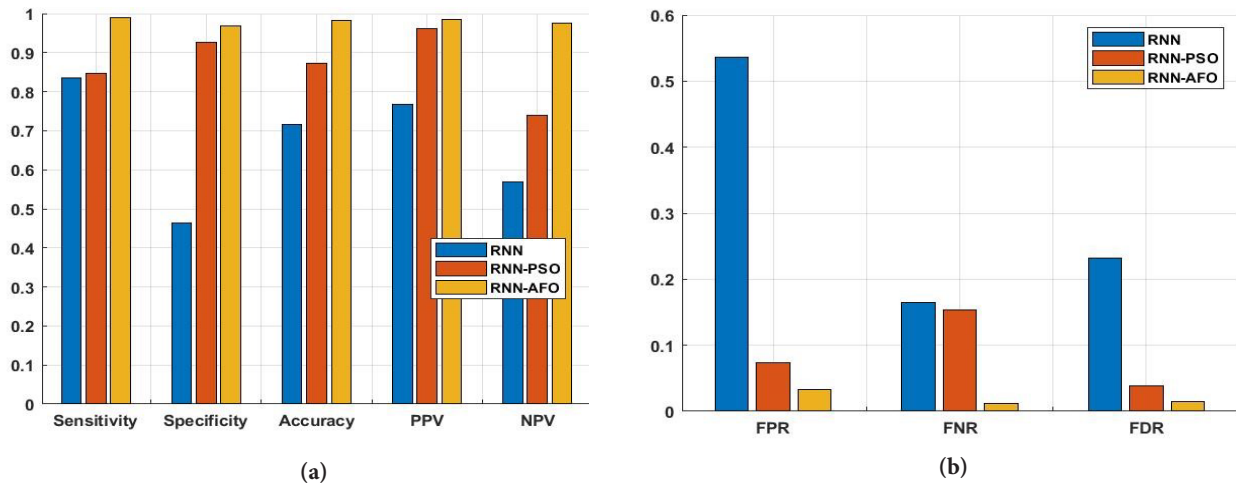


Figure 5 (a and b). The comparative analysis of different prediction models for kidney disease dataset.

Analysis of dataset related to heart disease

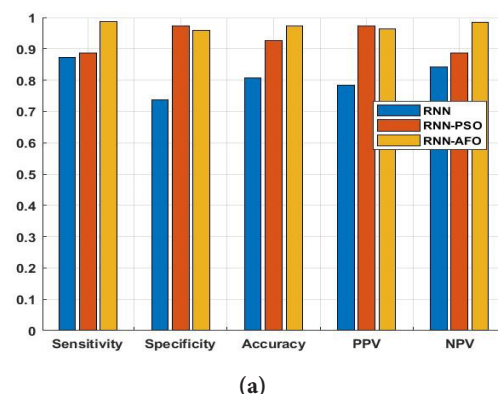
The suggested prediction model is evaluated and compared to other algorithms regarding accuracy, sensitivity, specificity,

PPV, NPV, FPR, FNR, and FDR. Table 3 depicts a study of performance indicators for heart disease datasets from diverse regions, including Cleveland, Hungary, and Switzerland.

Areas	Cleveland			Hungary			Switzerland		
Performance Metrics	RNN	RNN-PSO	RNN-AFO	RNN	RNN-PSO	RNN-AFO	RNN	RNN-PSO	RNN-AFO
Sensitivity	0.873418	0.886076	0.987342	0.414634	1	0.926829	0.5	0.5	1
Specificity	0.736111	0.972222	0.958333	1	0.915094	1	0.966102	0.983051	0.966102
Accuracy	0.807947	0.927152	0.97351	0.836735	0.938776	0.979592	0.95082	0.967213	0.967213
PPV	0.784091	0.972222	0.962963	1	0.82	1	0.333333	0.5	0.5
NPV	0.84127	0.886076	0.985714	0.815385	1	0.972477	0.982759	0.983051	1
FPR	0.263889	0.027778	0.041667	0	0.084906	0	0.033898	0.016949	0.033898
FNR	0.126582	0.113924	0.012658	0.585366	0	0.073171	0.5	0.5	0
FDR	0.215909	0.027778	0.037037	0	0.18	0	0.666667	0.5	0.5

Cleveland

The RNN-AFO model performs exceptionally well in Cleveland's heart disease prediction, as shown in Table 3 and Figures 6(a) and 6(b). It outperforms RNN and RNN-PSO with an accuracy of 0.97351 and 0.807947 and 0.927152, respectively. With a score of 0.987342, sensitivity the model can accurately detect most genuine positive instances. RNN-AFO's specificity is 0.958333, which is less than RNN-PSO's, but has great capacity to identify genuine negatives. Hence, RNN-AFO has greater prediction power for heart disease for Cleveland dataset.



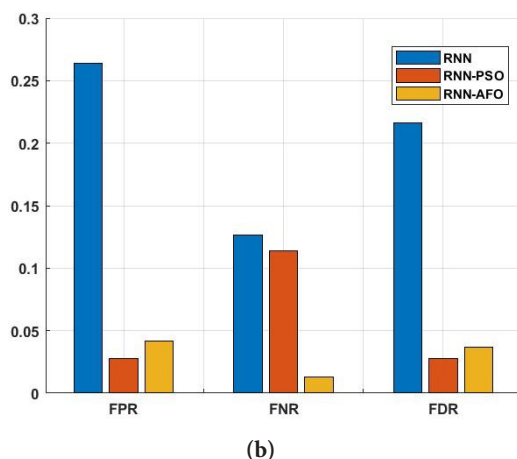


Figure 6 (a and b). The comparative analysis of different prediction models for the heart disease dataset of Cleveland.

Hungary

After examining the Hungary dataset from Table 3 and Figures 7(a) and 7(b), it is evident that the RNN-AFO model has much greater accuracy (0.979) than RNN (0.836) and RNN-PSO (0.938). With an ideal sensitivity score of 1.0, it indicates an extraordinary capacity to detect genuine positives. Specificity is 1.0, comparable to the RNN model, and no false positives were found—a significant improvement over the 0.915 of the RNN-PSO. Hence, RNN-AFO model efficiently predicts cardiac disease in the Hungarian dataset.

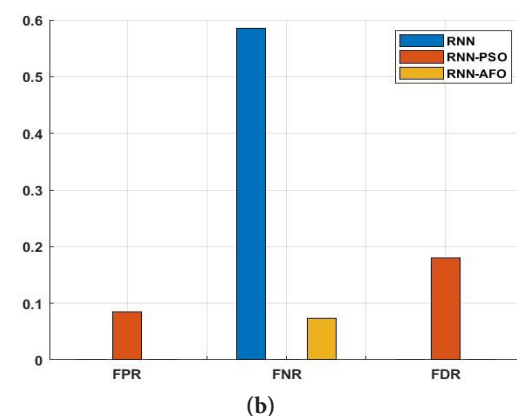
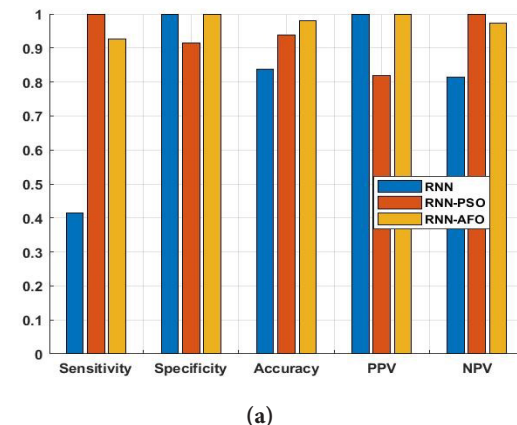


Figure 7 (a and b). The comparative analysis of different prediction models for the heart disease dataset of Hungary.

Switzerland

The accuracy of the RNN-AFO model for the Switzerland heart disease dataset is 0.967213 (Table 3 and Figures 8(a) and 8(b)). The existing models, which were less effective are not mentioned in the extract and are surpassed by this number. The perfect sensitivity (1.0), good specificity (0.966102), and high accuracy of the RNN-AFO model demonstrate its better prediction capabilities when used on the Switzerland dataset.

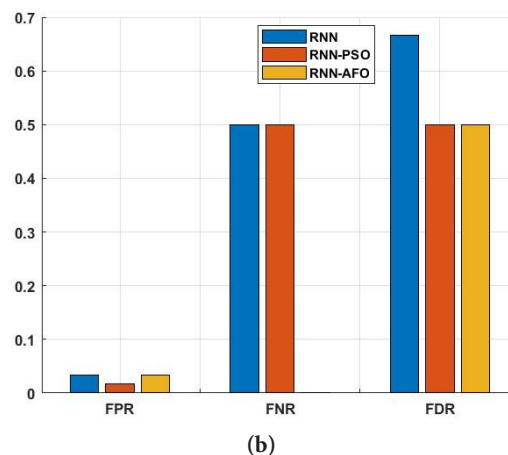
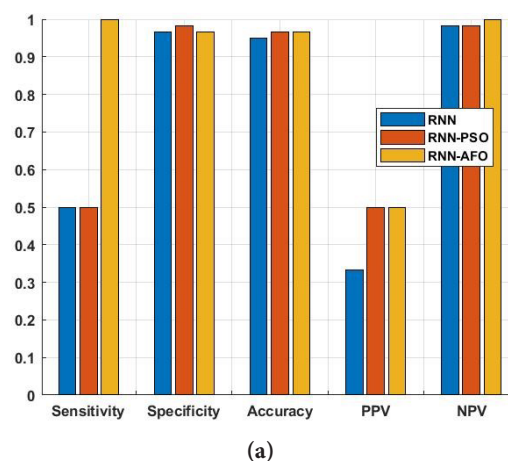


Figure 8 (a and b). The comparative analysis of different prediction models for the Heart disease dataset of Switzerland.

Conclusions

In conclusion, this study underscores the profound psychological impact of acne and the need for comprehensive care that addresses dermatological and psychological needs. By understanding the multifaceted nature of acne and its effects on mental health, healthcare providers can offer more effective and compassionate care, ultimately improving the quality of life for individuals affected by this common skin condition. This study developed a novel multi-disease prediction model utilizing RNN and AF optimization. Designed and tested on MATLAB, it shows improved performance over RNN and RNN-PSO models in accuracy, sensitivity, and specificity. It represents a breakthrough in healthcare monitoring systems, as evident by its use on a variety of datasets encompassing diabetes, heart disease, and kidney disease, which demonstrate its proficiency. Better patient care and outcomes may result from more prompt and efficient clinical setting interventions brought about by increased forecasting accuracy. The use of the AFA to optimize

RNNs creates new opportunities for using ML models in the medical field.

An important next step is to examine the model's scalability and adaptation to different illnesses and bigger datasets. Evaluating its performance in real clinical environments would be beneficial. Integrating the model with electronic health records could enhance predictive analytics in healthcare. Future research could explore integrating additional optimization methods with RNNs to enhance prediction accuracy. Additionally, comparing the model's performance with DL and AI techniques may reveal optimal approaches for analyzing medical data.

In summary, our study establishes a solid foundation for RNN-based multi-disease prediction models enhanced with AF optimization. Further research is needed to explore additional opportunities to advance AI in healthcare, potentially improving patient outcomes and administrative efficiency.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

1. Mosadeghrad AM. Factors affecting medical service quality. *Iran J Public Health*. 2014;43(2):210-220.
2. Sagahyroon A. Remote patients monitoring: challenges. In: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC). 2017:1-4. <https://doi.org/10.1109/CCWC.2017.7868460>
3. Chaudhury S, Paul D, Mukherjee R, Haldar S. Internet of thing based healthcare monitoring system. In: 2017 8th annual industrial automation and electromechanical engineering conference (IEMECON). 2017:346-349. <https://doi.org/10.1109/IEMECON.2017.8079620>
4. Aceto G, Persico V, Pescapé A. The role of information and communication technologies in healthcare: taxonomies, perspectives, and challenges. *J Netw Comput App*. 2018;107:125-154. <https://doi.org/10.1016/j.jnca.2018.02.008>
5. Manogaran G, Varadarajan R, Lopez D, Kumar PM, Sundarasekar R, Thota C. A new architecture of internet of things and big data ecosystem for secured smart healthcare monitoring and alerting system. *Future Gener Comput Syst*. 2018;82:375-387. <https://doi.org/10.1016/j.future.2017.10.045>
6. Kaur P, Kumar R, Kumar M. A healthcare monitoring system using random forest and internet of things (IoT). *Multimed Tools Appl*. 2019;78:19905-19916. <https://doi.org/10.1007/s11042-019-7327-8>
7. Kumar PM, Lokesh S, Varadarajan R, Babu GC, Parthasarathy P. Cloud and IoT based disease prediction and diagnosis system for healthcare using Fuzzy neural classifier. *Future Gener Comput Syst*. 2018;86:527-534. <https://doi.org/10.1016/j.future.2018.04.036>
8. Verma P, Sood SK. Cloud-centric IoT based disease diagnosis healthcare framework. *J Parallel Distrib Comput*. 2018;116:27-38. <https://doi.org/10.1016/j.jpdc.2017.11.018>
9. Ghorbani H, Izadyar M, Deilami HA, Ahmadzadegan MH. Massive DDOS occurrence investigation in future IoT devices. In: 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI). 2018:695-698. <https://doi.org/10.1109/ISRITI.2018.8864445>
10. Srinivas J, Das AK, Kumar N, Rodrigues JJ. Cloud centric authentication for wearable healthcare monitoring system. *IEEE Trans Dependable Secure Comput*. 2018;17(5):942-956. <https://doi.org/10.1109/TDSC.2018.2828306>
11. Amin MS, Chiam YK, Varathan KD. Identification of significant features and data mining techniques in predicting heart disease. *Telemat Inform*. 2019;36:82-93. <https://doi.org/10.1016/j.tele.2018.11.007>
12. Mathan K, Kumar PM, Panchatcharam P, Manogaran G, Varadarajan R. A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. *Des Autom Embed Syst*. 2018;22:225-242. <https://doi.org/10.1007/s10617-018-9205-4>
13. Reddy GT, Khare N. An efficient system for heart disease prediction using hybrid OFBAT with rule-based fuzzy logic model. *J Circuits Syst Comput*. 2017;26(04):1750061. <https://doi.org/10.1142/S021812661750061X>
14. Ahmad M, Tundjungsari V, Widiarti D, Amalia P, Rachmawati UA. Diagnostic decision support system of chronic kidney disease using support vector machine. In: 2017 second international conference on informatics and computing (ICIC). 2017:1-4. <https://doi.org/10.1109/IAC.2017.8280576>
15. Mshali H, Lemlouma T, Moloney M, Magoni D. A survey on health monitoring systems for health smart homes. *Int J Ind Ergon*. 2018;66:26-56. <https://doi.org/10.1016/j.ergon.2018.02.002>
16. Nilashi M, bin Ibrahim O, Ahmadi H, Shahmoradi L. An analytical method for diseases prediction using machine learning techniques. *Comput Chem Eng*. 2017;106:212-223. <https://doi.org/10.1016/j.compchemeng.2017.06.011>
17. Yuvaraj N, SriPreethaa KR. Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster. *Cluster Comput*. 2019;22:1-9. <https://doi.org/10.1007/s10586-017-1532-x>
18. Badawy M, Ramadan N, Hefny HA. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. *J Electr Syst Inf Technol*. 2023;10(1):40. <https://doi.org/10.1186/s43067-023-00108-y>
19. Wu Y, Zhang L, Bhatti UA, Huang M. Interpretable machine learning for personalized medical recommendations: a LIME-based approach. *Diagnostics (Basel)*. 2023;13(16):2681. <https://doi.org/10.3390/diagnostics13162681>
20. Duneja A, Puyalnithi T, Vankadara MV, Chilamkurti N. Analysis of inter-concept dependencies in disease diagnostic cognitive maps using recurrent neural network and genetic algorithms in time series clinical data for targeted treatment. *J Ambient Intell Humaniz Comput*. 2019;10:3915-3923. <https://doi.org/10.1007/s12652-018-1116-5>
21. Park DJ, Park MW, Lee H, Kim YJ, Kim Y, Park YH. Development of machine learning model for diagnostic disease prediction based on laboratory tests. *Sci Rep*. 2021;11(1):7567. <https://doi.org/10.1038/s41598-021-87171-5>
22. Vidhya K, Shanmugalakshmi R. Modified adaptive neuro-fuzzy inference system (M-ANFIS) based multi-disease analysis of healthcare big data. *J Supercomput*. 2020;76(11):8657-8678. <https://doi.org/10.1007/s11227-019-03132-w>
23. Harimoorthy K, Thangavelu M. Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. *J Ambient Intell Humaniz Comput*. 2021;12(3):3715-3723. <https://doi.org/10.1007/s12652-019-01652-0>
24. Ali F, El-Sappagh S, Islam SR, Kwak D, Ali A, Imran M, et al. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Inf Fusion*. 2020;63:208-222. <https://doi.org/10.1016/j.inffus.2020.06.008>
25. Chen M, Hao Y, Hwang K, Wang L, Wang L. Disease prediction by machine learning over big data from healthcare communities. *IEEE Access*. 2017;5:8869-88679. <https://doi.org/10.1109/access.2017.2694446>
26. Cheng L, Wu XH, Wang Y. Artificial flora (AF) optimization algorithm. *Appl Sci*. 2018;8(3):329. <https://doi.org/10.3390/app8030329>