A Quantum Mixed-State Self-Variational Onsager Neural Network with Leech Growth Algorithm-based Clinical Decision Support System for Accurate Heart Disease Detection

Fan Chenglin[1], Tai-Hoon Kim[2]*

[1,2]*Computer Engineering, Chonnam National University (CNU), South Korea.

[1]fanchenglin0625@jnu.ac.kr, [2]*taihoonn@chonnam.ac.kr
Corresponding author E-mail ID: taihoonn@chonnam.ac.kr

Abstract: Heart disease has been among the major causes of death in the world, and unless it is detected properly and on time, it cannot be treated effectively. Conventional methods of diagnosis generally have a problem with the complex clinical data, which results in misdiagnosis and delayed interventions, and that is why advanced computational methods are the ideal solution to this need. In this regard, the proposed research proposes the Quantum Mixed-State Self Variational Onsager Neural Network with Leech Growth Algorithm (QM-SSVONNet-LGA) to identify heart disease. At first, the Kaggle heart disease dataset is processed and prepared clinically according to the Z-score Min–Max Normalization (Z-M-MN) to provide consistency and reliability. QM-SSVONNet model represents intricate patterns in the data, whereas the Leech Growth Algorithm is an optimization of network parameters to achieve better convergence and accuracy. Experimental findings indicate that the proposed model works better than the existing approaches with a 99.1% accuracy rate, 98.7% precision, 99.5% specificity, and a low error rate of 0.9%. To summarize, QM-SSVONNet-LGA is a powerful, accurate, and effective model of reliable heart disease detection.

Keywords: Heart Disease Detection, Leech Growth Algorithm, Medical Data Preprocessing, Z-score Min–Max Normalization, Clinical Decision Support.

1. INTRODUCTION

Cardiovascular disease (CVD) has been identified as the most prevalent cause of mortality on earth, and the World Health Organization estimates 24.5 million deaths due to cardiac risk factors, such as hypertension, obesity, diabetes, and smoking, by the year 2030 (Almazroi et al., 2023). The intervention, as well as the early diagnosis, plays a vital role in preventing mortality since quick access to healthcare can save lives. The heart, as the most important organ, has to be carefully monitored and properly prognosticated (Al Reshan et al., 2023). The development of medical data has led to massive patient records and available clinical datasets. Using this knowledge along with the latest computational techniques, including machine learning (ML) and deep learning (DL), allows for analyzing, classifying,

and predicting heart diseases and conditions, as well as aiding physicians in providing quality and timely treatment (Nandy et al., 2023).

Diagnosis of heart disease is usually at its late stages, and therefore, treatment is not effective, and the risks of death may be fatal (Liang & Guo, 2023). Symptoms like irregular heartbeat, swelling of legs, rapid weight gain, and sleep disturbance are typical and non-specific, and they can often appear like other health complications, particularly among older patients (Ramesh & Lakshmanna, 2024). This resemblance makes it difficult to make a diagnosis early and correctly, resulting in a late intervention. Although there is a lot of clinical information available, it is difficult to draw useful insights that can be used to make predictions in due time (Khan Mamun & Elfouly, 2023). Traditional procedures are based on manual evaluation and subjective opinion and can be unequal. Thus, it is urgent to develop intelligent systems aimed at utilizing patient data and identifying cardiovascular risks beforehand, as well as helping health facilities to make correct clinical judgments (Kumar et al., 2023).

The increasing rates of incidence of heart disease and its high mortality rate point to the necessity of early detection systems. By processing detailed clinical data with the help of machine learning, it is possible to change the patient care by detecting the patterns of risks and predicting possible heart diseases. A predictive system will facilitate supporting patients with constant monitoring, faster diagnosis, and custom health recommendations, which will eventually save their lives. Moreover, these systems may help medical professionals to make evidence-based decisions effectively and allow their users to learn about their health. The paper suggests a model of deep learning to detect heart disease to enhance the accuracy of the diagnostics and preventive medical care.

Novelty and Contributions:

- The research presents a new neural network, Quantum Mixed-State Self Variational Onsager Neural Network (QM-SSVONNet), to predict heart diseases, as it combines quantum mixed-state representation with self-attention and Onsager dynamics to obtain the most important features.
- QM-SSVONNet uses the Leech Growth Algorithm (LGA) as its optimizer and offers a bio-inspired method to trade exploration and exploitation to improve convergence and accuracy.
- Z-M-MN is applied to preprocess heart disease data, providing strong and normalized input and making the model work better.
- The framework provides an integrated end-to-end flow and integrates data collection, preprocessing, model training, and optimization in a single framework to minimize the computational load.
- The model utilizes quantum-inspired representations to learn the complex nonlinear correlations in clinical information and enhance sensitivity and specificity in identifying heart diseases.

The relevant literature is carefully assessed in Section 2. Section 3 provides a detailed explanation of the methodologies used in this investigation. Section 4 discusses the outcome and its implications. Personal views and recommendations for additional research are included in Section 5.

2. LITERATURE SURVEY

Kaur et al., (2023) presented an automated method of identifying coronary heart diseases with the help of a Back Propagation Artificial Neural Network (BPANN). The architecture was tested to determine the dependencies of performance on the varied architecture of hidden layer neurons and the functions of performance. A tan-hyperbolic weighted function that has six hidden layers gave a better classification performance. Statistical validation and k-fold cross-validation were used to obtain high predictive performance and robustness. Nevertheless, deep architecture augmented the complexity of computation and reduced the interpretability.

Shrivastava et al., (2023) created a hybrid model to predict heart diseases based on Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (Bi-LSTM). It followed the procedure of data preprocessing to deal with missing and skewed records, after which feature selection was done with the help of an Extra Trees Classifier (ETC), and ultimate classification was carried out with the help of CNN-BiLSTM. The approach increased the diagnostic precision and learning. Nonetheless, deep architecture combined enhanced computational complexity and decreased interpretability to clinical decision support.

Manikandan et al., (2024) examined how machine learning classifiers (Logistic Regression (LR), Decision Tree (DT), and Support Vector Machine (SVM)) with Boruta feature selection can be used to predict heart disease. It was done by choosing the most pertinent clinical attributes to improve the overall performance of classification and decrease redundancy. Selecting features enhanced overall predictive efficiency, with LR having better accuracy. Nevertheless, it was based on traditional classifiers, and it did not have enough variety of features, which impeded the modeling of intricate nonlinear cardiac patterns.

Narasimhan & Victor, (2025) introduced a two-phase heart disease prediction architecture with the help of machine learning methods. The process involved the use of SelectKBest feature selection ranking on chi-square, mutual information, and F-statistic features, and then the use of Random Forest (RF), k-nearest neighbors (KNN), decision tree (DT), support vector machine (SVM), Naive Bayes (NB), Logistic Regression (LR), Gradient Boost, and neural networks. Better accuracy of diagnosis and fewer classification errors had been secured, but the model results were inconsistent across heterogeneous clinical records.

2.1 Problem Statement

Heart disease is usually diagnosed at its mature stages and thus becomes less effective in treatment and also risky in causing deaths. The symptoms, such as irregular heartbeat, swelling of legs, sudden weight gain, as well as sleeping disturbances, are not specific and can be confused with other conditions, and are thus difficult to diagnose early in any of the elderly patients. Regardless of the fact that there is a lot of clinical data, traditional approaches are based on a manual analysis and the judgment of experts, which may be inconsistent and time-consuming. Consequently, the necessity to design smart systems capable of processing patient data to forecast heart disease most efficiently and aid in making a timely clinical decision is particularly acute.

3. PROPOSED METHODOLOGY

This section introduces the proposed Quantum Mixed-State Self Variational Onsager Neural Networks with Leech Growth Algorithm (QM-SSVONNet-LGA) to detect heart disease. The data are first determined on the basis of the Kaggle heart disease dataset, which comprises clinically structured data that is linked to different heart diseases. Z-score Min -Max Normalization (Z-M-MN) is then applied to the gathered data to normalize the features and enhance the convergence of the model. It is then followed by the application of the QM-SSVONNet model, which is useful in identifying patterns of heart disease, and its loss minimization is performed by the Leech Growth Algorithm (LGA), which is more effective in improving exploration and exploitation to produce accurate and consistent outcomes. The proposed architecture is described in Figure 1.
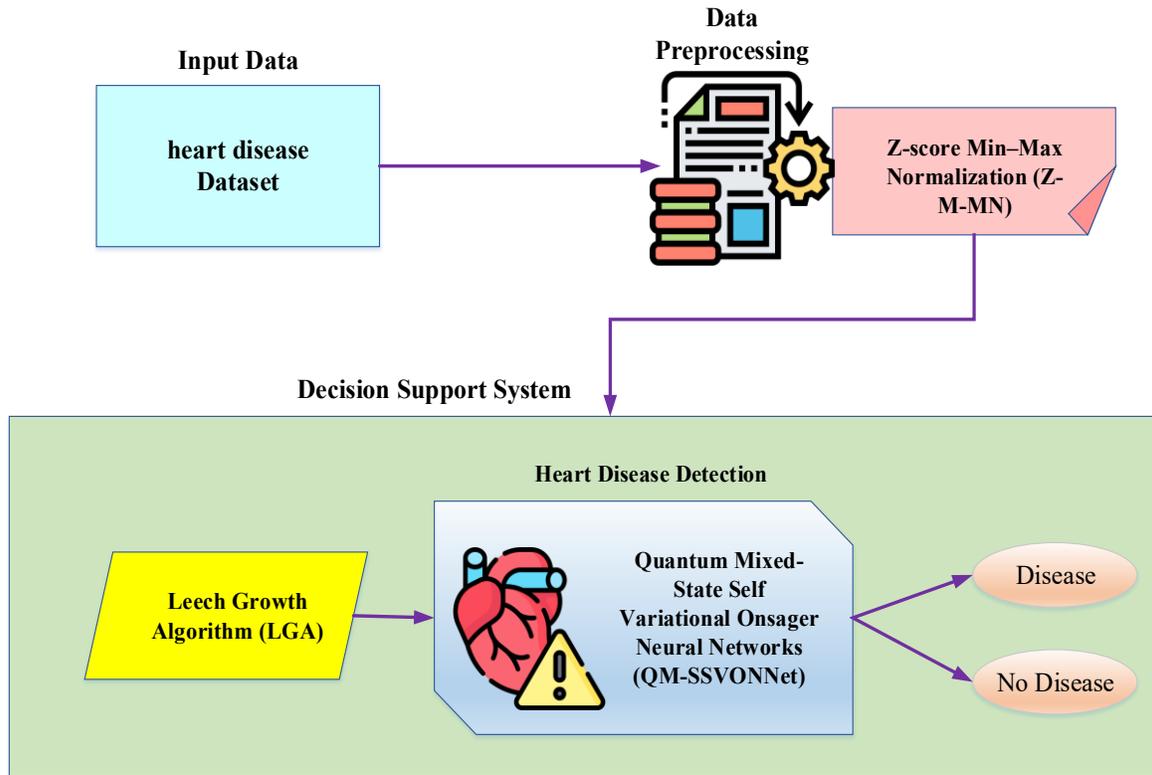
Figure 1. Workflow of the Proposed QM-SSVONNet-LGA Architecture.

3.1 Data Collection

The data regarding heart disease is sourced from the Kaggle dataset, which offers clinically structured data linked to heart problems. The dataset represents the demographic information of patients, observations, and diagnostic variables, which are combined to reflect the status of heart health. The data is arranged in a table form, making it suitable for data analysis. The data collected is then preprocessed before training and testing, ensuring higher quality and authenticity of the data regarding missing values, removing inconsistencies, smoothening out noise, and standardizing data features for accurate prediction of heart disease.

3.2 Data preprocessing using Z-score Min–Max Normalization (Z-M-MN)

Data preprocessing involves the application of the combined Z-score Min-Max Normalization (Z-M-MN) (Alshdaifat et al., 2021) approach. The strategy aims to enhance data quality and efficiency of learning. Firstly, the Min-Max Normalization method normalizes each dimension into a specific range, usually between 0 and 1, without affecting the relationships between the values of each dimension. This step involves applying each dimension's value using equation (1).

$$w' = \frac{w - \min_B}{\max_B - \min_B}\left(new\_\max_B - new\_\min_B\right) + new\_\min_B \qquad (1).$$

where, w' is the normalized value, w is the value of the original feature, $\min_B$ and $\max_B$ are the minimum and maximum values of the feature B, and $new\_\min_B$ and $new\_\max_B$ are the minimum and maximum values of the range to be targeted. Next, the Z-score normalization technique is used to mitigate the effect of outliers in the dataset by normalizing the features using statistical properties. The process is shown as equation (2).

$$w' = \frac{w - \mu}{\sigma} \qquad (2).$$

where w' is the standardized value, w is the original value of the feature, μ is the mean of the feature, and σ is the standard deviation of the feature. This step normalizes the original feature values to a distribution with a zero mean and a standard deviation of one. The normalized data are then passed to the classification phase.

3.3 Heart Disease Detection using Quantum Mixed-State Self Variational Onsager Neural Networks (QM-SSVONNet)

The task of heart disease diagnosis using the Quantum Mixed-State Self Variational Onsager Neural Networks (QM-SSVONNet) (Chen et al., 2025; Huang et al., 2022) is performed through the integration of quantum mixed-state self-attention and the use of the variational Onsager learning technique in understanding the various interactions involving cardiac symptoms. The clinical parameters of the patients are converted into quantum features, and similarity-driven attention focuses the network on the discriminating features.

The quantum mixed-state self-attention $\alpha_{t,k}$ calculates the similarity between the query and key representations with the following equation (3):

$$\alpha_{t,k} = us\left(\rho_{t,r}\sigma_{k,p}\right) \qquad (3).$$

Where, $\rho_{t,r}$ denotes the density matrix of mixed state of the tth query feature vector obtained from patients, $\sigma_{k,p}$ denotes the density matrix of mixed state of the kth key feature vector, us denotes the trace operator that measures the overlap of a quantum state. Attention Weights are normalized as follows using equation (4).

$$\widetilde{\alpha}_{t,k} = \frac{\alpha_{t,k}}{\sum_{n=1}^{T} \alpha_{t,n}} \tag{4}$$

Where, T depicts the total number of input features, $\alpha_{t,n}$ symbolizes the similarity between the tth query and the nth key, while $\widetilde{\alpha}_{t,k}$ depicts a probability distribution that emphasizes the role of each feature in the determination of heart diseases. The variational Onsager neural network corresponds to the dynamic process of the cardiac system. The free energy density is given by equation (5).

$$g(a) = g^* \left[\widetilde{g}(a) - \widetilde{g}(0) \right] \tag{5}$$

where, a represents the normalized state variables, $\widetilde{g}$ represents the approximation of the free energy using the neural network, and g^* is a normalization constant. On the other hand, the dissipation potential can be written as given in equation (6).

$$\psi(a,x) = \psi^* \left[\widetilde{\psi}(a,x) - \widetilde{\psi}(a,x) - \frac{\partial \widetilde{\psi}}{\partial x} \Big|_{x=0} \bullet x \right] \tag{6}$$

Here, the $\widetilde{\psi}$ function tries to model the dissipation part and ψ^* acts as a characteristic scale, which imposes the convexity constraint on the process variable x. The learning task aims to maximize the objective given by the equation (7).

$$\ell = \alpha_{PDE} \ell_{PDE} + \alpha_{BC} \ell_{BC} \tag{7}$$

Where ℓ_{PDE} denotes the measurement of the residual of the governing equations of the heart, ℓ_{BC} denotes the measurement of boundary condition satisfaction, and $\alpha_{PDE}, \alpha_{BC}$ are coefficients used to balance the impact of each term to obtain correct and accurate heart disease classification. The QM-SSVONNet loss function is then optimized with the Leech Growth Algorithm.

3.4 QM-SSVONNet optimizer using Leech Growth Algorithm (LGA)

QM-SSVONNet adjusts its loss function ℓ with the help of the Leech Growth Algorithm (LGA) (Wu et al., 2025), which is based on some natural actions like roaming randomly, foraging, nesting, and mating, displayed by leeches. Here, every leech corresponds to a solution in the solution space, and its current position is updated in every iteration. The maturation factor $N(t)$ facilitates the process of exploring and exploiting in LGA. While exploring, leeches move randomly and look for safer locations. On exploiting,

they nest and mate to reach globally optimal locations. Environmental conditions, mating skills, and greedy choices help in achieving strong convergence in LGA. LGA improves the robustness and efficacy of QM-SSVONNet in identifying heart diseases. The LGA process is described in Algorithm 1.

Algorithm 1: LGA

Initialize the leech population L_{pop} with M individuals, each with a position L_j

Set maximum iterations T, problem dimension E

for t = 1 to T do

  Update maturation factor $N(t)$ based on current iteration

  for each leech L_j in L_{pop} do

    if $|N(t)| >= 0.5|$ then   *// Exploration phase*

      Choose random movement or foraging.

      Update L_j toward safer positions based on population and environment.

    else   *// Exploitation phase*

      Choose nesting or mating.

      Move L_j toward the global best and average population positions.

    end if

    Evaluate fitness $f_j = \ell(L_j)$

  end for

  Perform greedy selection: retain better positions for next iteration.

end for

Return L_{best} minimizing QM-SSVONNet loss ℓ

---

4. RESULTS

The section provides the experimental analysis of the proposed QM-SSVONNet-LGA to detect heart disease. The model is coded in Python 3.9 and is based on TensorFlow and Keras, and data processing and performance evaluation are performed using NumPy, Pandas, and Scikit-learn. The experiments are performed on a system that has an Intel Core i7 (2.8 GHz), 16 GB RAM, 4 GB graphics card, and Windows 64-bit. Training and testing are carried out using the Kaggle Heart Disease. The key parameters of the simulation used in the experiments are enumerated in Table 1.

Table 1. Simulation Parameters.

Parameters	Description
Operating System	Windows 64-bit
Dataset	Kaggle Heart Disease

Data Type	Structured Clinical Data
Proposed Model	QM-SSVONNet
Optimizer	Leech Growth Algorithm (LGA)
Number of Epochs	100

4.1 Dataset Description

The Heart Disease dataset (https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset), which was created in 1988, is a collection of four databases (Cleveland, Hungary, Switzerland, and Long Beach V). It also has a total of 76 attributes, among which it has a target attribute, but the most common studies usually consider 14 key attributes that are important in the prediction of heart disease. The target attribute signifies the existence of heart disease, with 0 being the absence of the disease and 1 being the presence of the disease. To evaluate the model, the dataset will be divided into training and testing constituents, where 80%of the data is used in the training model and the remaining 20% is used in testing the model performance.

4.2 Performance Evaluation

The proposed QM-SSVONNet-LGA model is applied to the Kaggle heart disease dataset to compare it with the existing models, such as BPANN (Kaur et al., 2023), CNN-BiLSTM (Shrivastava et al., 2023), LR (Manikandan et al., 2024), and KNN (Narasimhan & Victor, 2025). The most important metrics of evaluation of the performance include accuracy, precision, specificity, recall, F1-score, and error rate. The suggested QM-SSVONNet-LGA has shown better results in all of the metrics, which means that it is strong enough to properly identify the cases of heart diseases, and it has better predictive performance and reliability compared to both traditional machine learning and deep learning-based versions.
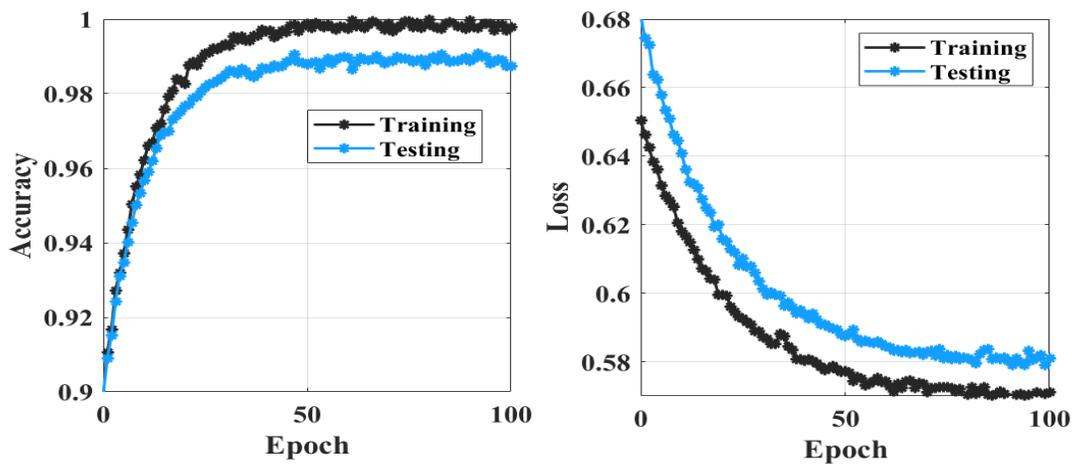

Figure 2. Training and Testing Performance of QM-SSVONNet-LGA.

Figure 2 (Plot 1) depicts the accuracy evolution of the proposed QM-SSVONNet-LGA in 100 epochs of the training and testing phases. Training accuracy gradually rises to almost 0.99, whereas testing accuracy follows closely with a steady value of around 0.99, which is high model generalization. The second plot shows loss curves of training and testing sets. The decrease of both losses is constant with epochs; both training and testing losses are about 0.58 and 0.60, respectively, which represents good learning and low overfitting.
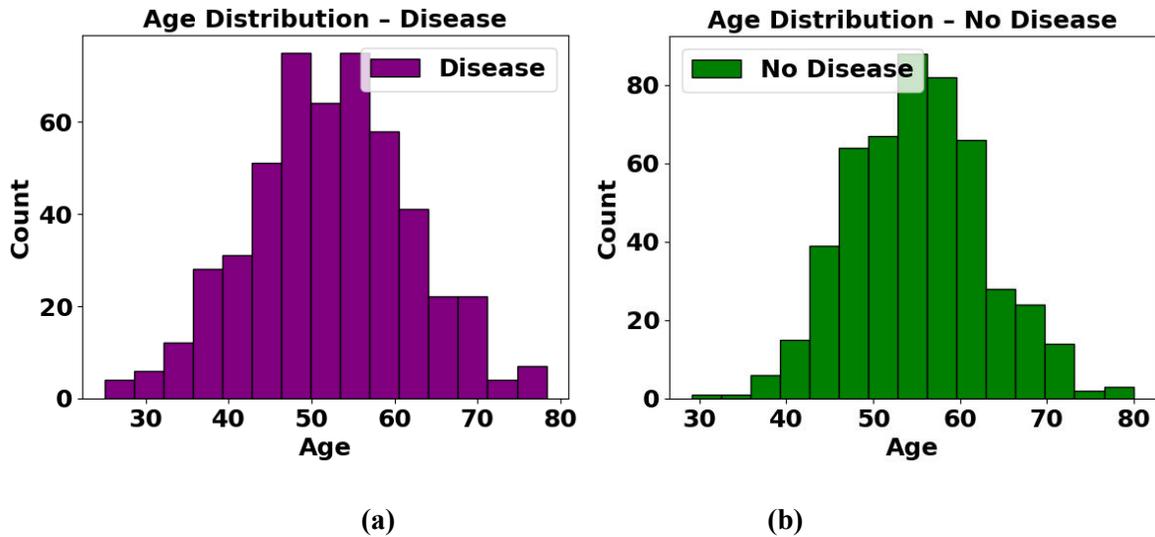

(a) **(b)**

Figure 3. Age Distribution of Patients with (a) Heart Disease, (b) No Disease.

Figure 3(a) depicts the distribution of the age of patients affected by heart disease, where the frequency of the estimation is highest in the age group, which means the middle-aged individuals are at risk of heart disease. Figure 3(b) shows the age distribution of patients not having heart disease, with the highest concentration at 55, which indicates a relatively balanced distribution and the inclusion of age as an important predictive factor of heart disease.

Table 2. Comparative Performance of Heart Disease Detection Models.

Model	Accuracy (%)	Precision (%)	Specificity (%)	Recall (%)	F1-Score (%)	Error Rate (%)
BPANN (Kaur et al., 2023)	85.2	84.1	86.5	83.0	83.5	14.8
CNN–BiLSTM (Shrivastava et al., 2023)	88.5	87.2	89.0	86.5	86.8	11.5
LR (Manikandan et	81.7	80.5	83.2	79.8	80.1	18.3

al., 2024)						
KNN (Narasimhan & Victor, 2025)	83.9	82.7	85.1	82.0	82.3	16.1
Proposed QM-SSVONNet-LGA	99.1	98.7	99.5	98.9	98.8	0.9

A comparative analysis of different heart disease detection models in terms of key performance metrics is provided in Table 2. The conventional approaches, including BPANN, CNN-BiLSTM, LR, and KNN, demonstrate the middle range of accuracy of 81.7-88.5 with the relevant measure of precision, recall, and F1-score that demonstrate poor predictive ability. Conversely, the proposed QM-SSVONNet-LGA model exhibits better performance in all metrics and a 99.1% accuracy, 98.7% precision, 98.9% recall, and 98.8% F1-score with a minimum error rate of 0.9%. These findings demonstrate the usefulness of the proposed model in precise and dependable heart disease diagnosis.

5. CONCLUSION

A new Quantum Mixed-State Self Variational Onsager Neural Network with Leech Growth Algorithm (QM-SSVONNet-LGA) is introduced in this study to detect heart disease in the most accurate way. As shown in the experimental results, the model has a better performance with an accuracy of 99.1%, a precision of 98.7%, a specificity of 99.5%, a recall of 98.9%, an F1-score of 98.8%, and a low error rate of 0.9% compared to current methods. There are complex trends that are observed in the clinical data, and the proposed method is effective enough to capture such trends and allow reliable and consistent predictions. The benefits of QM-SSVONNet-LGA are that it has a high predictive accuracy, strong generalization, and the Leech Growth Algorithm is an optimizer that has fast convergence, which makes it applicable in the real world of healthcare. The model, however, has higher computational and training time needs, and its performance might not match the various datasets, which will need additional validation. Future research could be aimed at combining multimodal clinical information, improving computational performance, and making it usable in real-time in clinical decision support systems to provide early and accurate diagnosis.

REFERENCE

[1]   Almazroi, A. A., Aldhahri, E. A., Bashir, S., & Ashfaq, S. (2023). A clinical decision support system for heart disease prediction using deep learning. *IEEE Access*, *11*, 61646-61659.

[2]   Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., & Shaikh, A. (2023). A robust heart disease prediction system using hybrid deep neural networks. *IEEE Access*, *11*, 121574-121591.

[3] Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V. G., Li, X., & Zakarya, M. (2023). An intelligent heart disease prediction system based on swarm-artificial neural network. *Neural Computing and Applications*, *35*(20), 14723-14737.

[4] Liang, Y., & Guo, C. (2023). Heart failure disease prediction and stratification with temporal electronic health records data using patient representation. *Biocybernetics and Biomedical Engineering*, *43*(1), 124-141.

[5] Ramesh, B., & Lakshmanna, K. (2024). A novel early detection and prevention of coronary heart disease framework using hybrid deep learning model and neural fuzzy inference system. *IEEe Access*, *12*, 26683-26695.

[6] Khan Mamun, M. M. R., & Elfouly, T. (2023). Detection of cardiovascular disease from clinical parameters using a one-dimensional convolutional neural network. *Bioengineering*, *10*(7), 796.

[7] Kumar, A., Singh, K. U., & Kumar, M. (2023). A clinical data analysis based diagnostic systems for heart disease prediction using ensemble method. *Big Data Mining and Analytics*, *6*(4), 513-525.

[8] Kaur, J., Khehra, B. S., & Singh, A. (2023). Back propagation artificial neural network for diagnose of the heart disease. *Journal of Reliable Intelligent Environments*, *9*(1), 57-85.

[9] Shrivastava, P. K., Sharma, M., & Kumar, A. (2023). HCBiLSTM: A hybrid model for predicting heart disease using CNN and BiLSTM algorithms. *Measurement: Sensors*, *25*, 100657.

[10] Manikandan, G., Pragadeesh, B., Manojkumar, V., Karthikeyan, A. L., Manikandan, R., & Gandomi, A. H. (2024). Classification models combined with Boruta feature selection for heart disease prediction. *Informatics in Medicine Unlocked*, *44*, 101442.

[11] Narasimhan, G., & Victor, A. (2025). Empirical analysis of predicting heart disease using diverse datasets and classification procedures of machine learning. *Ain Shams Engineering Journal*, *16*(8), 103470.

[12] Alshdaifat, E. A., Alshdaifat, D. A., Alsarhan, A., Hussein, F., & El-Salhi, S. M. D. F. S. (2021). The effect of preprocessing techniques, applied to numeric features, on classification algorithms' performance. *Data*, *6*(2), 11.

[13] Chen, F., Zhao, Q., Feng, L., Chen, C., Lin, Y., & Lin, J. (2025). Quantum mixed-state self-attention network. *Neural Networks*, *185*, 107123.

[14] Huang, S., He, Z., & Reina, C. (2022). Variational Onsager Neural Networks (VONNs): A thermodynamics-based variational learning strategy for non-equilibrium PDEs. *Journal of the Mechanics and Physics of Solids*, *163*, 104856.

[15] Wu, J., Su, Z., & Luo, W. (2025). Leech Growth Algorithm: a new meta-heuristic algorithm. *The Journal of Supercomputing*, *81*(8), 988.

[16] https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset