Received: 17 July 2023 | Revised: 3 October 2023 | Accepted: 6 November 2023

IET Quantum Communication

DOI: 10.1049/qtc2.12078

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Quantum intelligence in medicine: Empowering thyroid disease prediction through advanced machine learning

Mohemmed Sha

Department of Software Engineering, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

Correspondence
Mohemmed Sha.
Email: ms.mohamed@psau.edu.sa

Funding information
Prince Sattam bin Abdulaziz University project number, Grant/Award Number: PSAU/2023/R/1445

Abstract

The medical information system is rich in datasets, but no intelligent systems can easily analyse the disease. Recently, ML (Machine Learning)-based algorithms have acted as a handy diagnostic tool to identify whether a person is affected by thyroid or not. However, they produced classification with low accuracy and led to misclassification. Hence, the proposed system combines quantum computing with ML techniques to enhance computational power and precision. The system employs modified QPSO (Quantum Particle Swarm Optimisation) for feature selection since its searching performance is better than that of conventional PSO for selecting the optimum global position of the particle, thus selecting the relevant feature. Whereas, the QSVM (Quantum Support Vector Machine) is implemented for more accurate classification than classical SVM, as it tends to capture complex patterns in data produced due to high dimensional feature space applied by quantum kernel functions. This combination of modified QPSO and QSVM tends to increase the performance accuracy significantly. The efficiency of the proposed model is measured based on derivative parameters, such as F-1-score, recall, precision and accuracy, with corresponding confusion matrix and ROC. Further, the classification is compared with other traditional approaches to predict the accuracy of the proposed model with traditional methods.

KEYWORDS
learning (artificial intelligence), quantum communication

1 | INTRODUCTION

Thyroid disease is a prevalent disease occurring worldwide, mostly caused by a deficiency of iodine in the body. The thyroid-based hormone is secreted and generated by the thyroid gland, which normalises the body's metabolism. The complicated stages of thyroid result in increased cholesterol, blood pressure and depression. Misclassification results in healthy patients enduring unnecessary treatments. The traditional methods include appropriate medical examination with several blood specimens for taking blood tests, which measure TSH (Thyroid-Stimulating Hormone), T3(Triiodothyronine) and T4(Thyroxin) levels. Thus, there is a requirement for a system to detect thyroid-based diseases in the initial stage of enlargement. In the case of the medical sector, ML algorithms play a significant part in thyroid disease analysis since it produces several classification methods, depending on which a system can be trained with a properly trained dataset of thyroid-affected patients and tend to identify and produce outcomes precisely. Based on advancements in technologies, such as parallel computing, big data, data mining and image and video processing, the prediction of thyroid diseases lacks accurate performance. Thus, QC (Quantum Computing)-based [1–3] classification [3–5] enhances the probability of predicting the disease [6]. Due to parallel quantum computing, a derivative of quantum state entanglement and superposition, QC-based algorithms [7, 8] are better than classical image processing approaches [9, 10].

Contemporarily, the ML techniques, such as RF (Random Forest) and SVM (Support Vector Machine) [11], have been deployed in the intimated paper to predict thyroid disorder. The experimental analysis has been performed using Python on the Spyder interface. The dataset based on UCI respiratory has been taken into account, which has been represented as the benchmark dataset for the classification of disorders. It consists of nearly 7200 illustrations along with 21 input traits and one output trait. Then, the pre-processing method has been processed for converting definite variables to numerical values and predicting the missing values in the dataset using the Min-Max normalisation technique. From the analysis, the SVM has produced better results with an accuracy of 93% and RF with 92%. A CAD (Computer Aided Diagnostic) system has been involved in categorising the thyroid nodules. Nearly 447 US images are used to predict and classify thyroid nodules. Here the features selection method is based on multi-objective PSO [12–14] to select non-redundant and relevant individuals. Then by employing RF and SVM, the classification is performed. Both optimisation and reduction techniques on attributes are deployed for the classification of thyroid nodules to improve precision. The suggested system produced an accuracy of 93.24%.

Similarly, based on the optimisation technique, a Graph Clustering ACO (Ant Colony Optimisation) [15] with an extreme ML algorithm is used to produce an efficient prediction of thyroid nodules. The optimal features [16] are selected for classification, which helps in the improvement of the classifier function. The main approach of the system is to classify benign along with malignant nodules from the dataset. The efficiency of feature selection, as well as classification, is measured based on specificity, accuracy, sensitivity and AUC (Area under Curve). Thus the intimated system produced better accuracy than univariate methods.

To solve the issue of binary classification of thyroid disease, efficient QC-based ML approaches are implemented in the proposed system. The dataset used in the respective method is the thyroid disease dataset. Then, the loaded data is pre-processed to eliminate the noise and handle missing values, thus enhancing the quality of the image. QPSO performs the optimisation of feature selection with quantum gates such as Hadamard and Pauli matrix, which tend to increase the feature selection method for better classification. The fitness function of selected features is measured to evaluate the over-fitting issues. Then, the selected feature is fed to train and test split, and the trained data is fed into the model to test its performance in classification. QSVM performs the classification of data, as it tends to achieve faster training time than classical SVM because of improved computation of QC. This method is specifically used in large datasets where training time is a major issue. It also enables the use of kernels that are hard to compute classically. Then the efficiency of the system is compared along other traditional classification algorithms, such as RF (Random Forest), K-NN (K-Nearest Neighbour) and SVM (Support Vector Machine). Further, the trained model is validated with test data in the prediction phase, and the efficacy of the system is analysed based on derivative parameters, such as recall, F1-score, accuracy and precision.

The main contribution of the study is as follows:

- To implement feature selection, by using modified QPSO for selecting the global best attribute suitable for classification.
- To employ QSVM to improve the classification of thyroid and non-thyroid based on the dataset.
- To compare the proposed model with traditional KNN, RF and SVM, to evaluate the efficiency of the system, and to analyse the performance of the system based on derivatives, such as F1-score, recall, precision and accuracy.

1.1 | Paper organisation

Based on the objectives, the present study is developed in a manner that deliberates the several techniques involved in thyroid disease prediction using ML algorithms. The classification method for predicting thyroid-affected and non-thyroid persons based on the given dataset is mentioned. Then Section 2 reviews and explores the various classifiers involved in enhanced classification. Following this, the step-wise process involved in classification and comparison with traditional methods is presented in Section 3. Consequently, the experimental results and their corresponding performance metrics are discussed in Section 4. Thus the overall study is concluded with the outcomes and future work of the proposed system in Section 5.

2 | REVIEW OF EXISTING WORK

Prediction and classification of non-thyroid and thyroid people are interpreted by several [17, 18] ML-based methods. These techniques explored by several researchers are reviewed and studied. Some significant problems identified from the existing methods are also portrayed in this section.

Some minor problems created in the thyroid gland can cause severe issues all over the body. Thyroid inefficiency leads to diseases such as hyperthyroidism and hypothyroidism. ML-based [19, 20] algorithms, namely, RF, ANN, DT (Decision Tree) and KNN [21], have been employed in the system to classify the medical images as present in the dataset for early prediction. Specificity and sensitivity have been implemented to represent whether the disease is present or not. From the analysis of applying ML techniques, it has been found that KNN has produced the least performance and projected 91% and 59% of specificity and sensitivity, respectively. Whereas, RF has produced a better classification rate of 91% and 94.8% of specificity and sensitivity, respectively, based on the dataset. Though, traditional system accomplished better performance in the classification, it lacks in factor such as accuracy. The model has been trained by using classification techniques, such as SVM, NB (Naïve Bayes), KNN and DT [22] to detect thyroid disease. A clinical dataset has been used, which consists of 807 patients with 583 females and 224 males. It has been found that from the dataset, 36 possess hyperthyroidism, 218 have hypothyroidism and 553 tend to be normal. The attributes deployed are gender, T3, TSH, T4, age and other classification features. The efficiency of each classifier has been evaluated based on accuracy. From the results, the DT has produced better classification when compared with other classifiers. The data of the traditional system is taken from the lab of Kashmir. The effective performance of the traditional system signified the result with better accuracy. The suggested paper has used ML algorithms, such as RF, DT, NB, multi-class classification and ANN to identify the occurrence of hypothyroidism present in the dataset. The model [23] has used pre-processing and feature selection for enhanced disease classification. In pre-processing technique, the data integration and data cleaning methods have been involved to obliterate the missing values and noise present in the dataset. Whereas, the feature selection method has been applied to identify the important attributes related to envisaging the output. The redundant data has been removed to optimise the precision, accuracy and recall and reduce computation time. Thus, the RF and DT [24] have produced better results than other comparative classifications. The efficiency of the existing model is signified through the results. However, it used DL-based technique, which requires a large set of data for the classification. The analysis of several classifiers has been applied in the intimated paper [25, 26] to predict the effective classifiers in thyroid disease prediction. Some of the ML-based algorithms implied in the model are logistic regression, KNN, SVM, NB and DT. The dataset which has been taken into consideration are data collected from DHQ teaching hospital based on three features, such as blood pressure, pulse rate and BMI (Body Mass Index), which directly associate with thyroid deficiency prediction. From the experimental results, the traditional research accomplished effective performance in the classification of thyroid disease with better accuracy. Besides, the multi class classification represents the efficiency of the existing model. The model has been categorised into three phases, in which the first phase has not used feature selection [27]. But, L1 and L2-based feature selection method has been applied in phases 2 and 3. From the analysis, it has been found that logistic regression and NB have produced better outcomes, with KNN producing an error rate of 2.16%. The conventional model efficiency is represented through the outcome of the classification with higher accuracy. However, it lacks in effective feature selection process and the missing data in the dataset. Similarly, DT is most commonly used to predict decisions for many large and complex datasets. The intimated paper [28] has used an ensemble classification method to determine hyperthyroidism and hypothyroidism in hormones. Hoeffding, RF and J48 have been used for classification. The real-time dataset has been applied to 499 thyroid patients. It has been used to detect the presence or absence of thyroid in the hormone based on TSH, T3 and T4 values. Based on the accuracy, batch size, number of folds, confidential tree factor, seed, time and the minimum number has been compared with several thyroid illnesses. The classification has been performed on all thyroid datasets with various attributes and iterations and equating the tree ensemble system. The outcome of the classification signifies the effective performance of the traditional method with higher accuracy.

Thyroid disorders are predicted by their occurrence and functional behaviour and are caused by the inadequate presence of thyroid hormones. Likewise, thyroid disease prediction in women based on CART (Classification and Regression Tree), RF and DT has been implemented in the model [29]. The dataset possesses nearly 3710 illustrations with 30 attributes. The decision variables have been separated into two classes such as positive as well as negative. The constant values are altered into numerical values with 1 as true and 0 as false. With the outcomes produced from the classifiers, a bagging ensemble method has been applied to enhance the performance of the model. This ensemble method employs prediction in a simultaneous process with training and testing data. From the results, the RF has outperformed other techniques. From the experimental outcome, the exiting model accomplished better performance with higher accuracy. Nearly 5 ML algorithms, namely, Boosting, NB, Bagging, SVM [30, 31] and J48, have been implied in the model [32] to compare the efficiency of the classifiers to detect thyroid disease. The datasets included in the classification are both serological and pathological annotations of the patients, with 21 features taken from the hospital. The dataset has been separated into 10 subsets of which nine subsets are used in training, and the remaining one subset is used in testing. The better performance of the traditional system has been represented through the results with the higher accuracy. However, in the case of the UCI dataset, several missing values are identified, which tends to lessen the efficiency of the classifiers.

Finally, the total mean of classification precision has been calculated. Here the test datasets have been measured independently, thus increasing the accuracy of the outcomes. As a result, the bagging classifier has performed better than other methods. The diagnosis accuracy has been predicted by implementing an ML algorithm, namely MMLP (Multiple Multi-Layer Perceptron) [33, 34]. The dataset used in the suggested system is the thyroid disease dataset processed from the UCI repository. It consists of 7200 patients, including 6 continuous features and 15 binary features deprived of any missing values. Based on Pearson's correlation, features have been selected which are considered significant risk factors in thyroid disease. Feature scaling and standardisation have been involved in the classification method to rescale the values of attributes starting from dynamic to a specific range. It has produced better accuracy and generalisation with reduced over-fitting issues.

ML and data mining techniques are widely used for thyroid disease prediction. The considered paper [35] has used PCA (Principal Component Analysis) and dimension reduction methods to reduce the input data dimension to classifiers for thyroid disease prediction. And data augmentation has also been implied to produce efficient data for the model. The real-time dataset has been used, and the experiments have been performed in the distributed background. Nearly, 3152 cases are present in the thyroid disease dataset, with 23 attributes have been used in the system. Data pre-processing has been performed to predict the invalid data and missing data and normalisation for standardisation of data. The classes are divided into class 0 (non-thyroid) and class 1 (thyroid). The standard deviation and mean for each class have been evaluated. The efficacy of the traditional model has been represented through the outcome of the experiment with higher accuracy. Besides, the experiment with the diverse experiment has been considered as the effective performance of the existing method. Likewise, the XGBoost method [36] has been deployed in the prediction of thyroid disease, and suitable attributes have been selected based on the XGBoost function. The dataset is divided based on normal, hypothyroid and hyperthyroid. Here the dataset has also been categorised into three sections such as training, testing as well as validation data. Then the training set has been applied on XGBoost, and the test data is tested with the pre-trained model. The effective performance of the conventional model has been signified through the results with better accuracy. However, the exiting model requires more attributes for the detection of thyroid disease.

Thus the model predicts the output with better accuracy by comparing the efficiency with KNN, DT and logistic regression. The intimated study [37] has used a hybrid optimisation algorithm, associating HFBO-RT2FSVM (Hybrid Fire-fly Butterfly Optimisation-Rough Type-2 Fuzzy SVM) based on feature selection. It has been used to classify thyroid diseases based on sensitivity, accuracy and specificity. The thyroid dataset consists of 29 features, which are pre-processed by denoising and predicting missing values. The outcome of the experiment signifies the effective performance of the conventional model with the higher accuracy. Here, the butterfly algorithm [38, 39] and firefly algorithms are joined together to form the HFBO approach. This approach helps in feature reduction and decreases the number of redundant and irrelevant features. As a result, the model has produced better accuracy by comparing with MKSVM (Mixed Kernel SVM) and IGWO Linear SVM (Improved Grey Wolf Optimisation Linear SVM). The efficiency of the traditional method is represented through the results with the higher accuracy of 0.943.

2.1 | Problem identification

Though different hybrid methods of ML algorithms are implemented in the prediction of thyroid disease, there are some complications in conventional methods which has to be overcome. Different issues are analysed and listed in this section.

- The suggested study used RF and SVM algorithms for classification, but it employed only limited parameters and automatic extraction of ROI in ultrasound image samples [12].
- The study employed XGBoost for classification and lacked in implementing the model based on real-time datasets, and used a large number of attributes to detect thyroid disease [36].
- ML algorithms based on fuzzy c-means clustering have been used in the model and have resulted in lower performance in categorising and decision patterns, which are used in providing effective treatment [40].

3 | PROPOSED METHODOLOGY

Thyroid disease is considered a multifarious and persistent contagion, produced due to the high level of hormones secreted in the thyroid gland. Several approaches are involved in the identification of thyroid diseases, such as blood results from the laboratory, data mining and ML techniques. However, the traditional methods have produced lower classification accuracy. So, optimised accuracy in the classification of thyroid-affected and non-thyroid-affected persons is performed by using quantum-based ML techniques. The proposed system implements modified QPSO for feature selection and QSVM for classification. The dataset used in the model is the thyroid disease dataset, consisting of 7200 instances with 29 attributes. The dataset consists of 2800 training sets and 972 testing sets. Initially, the medical data is transformed into quantum information. The conversion of medical data into the quantum data includes encoding the data into qubits. The methods of conversion includes quantum circuits, quantum programming language, quantum algorithms and quantum gates. Initially, in the conversion, every feature in the data is mapped to qubits. The technique of quantum encoding includes amplitude encoding which involves amplitude of qubits to the feature value. The process of quantum encoding makes the superposition, where qubits can signify numerous feature states. The qubits can be entangled, generating relationship among the states. The attributes used in the conversion of medical data into qubits includes quantum feature maps, quantum gates, quantum kernels and quantum circuits. The quantum feature maps convert the traditional feature representation into the quantum states, where the quantum circuit processes the operation on the qubits to convert the encoded data. Similarly, quantum circuits and gates process phase shifts and rotations on the qubits where the quantum kernels calculate the inner product among the quantum encoded data points. Likewise, the quantum vibrational circuit regulate the quantum states based on the method of optimisation.

Primarily, in the proposed system, the dataset is first loaded and generated for pre-processing method. The pre-processing is performed, which involves cleaning and deletion of unnecessary entries and columns. Processing the missing values and eliminating unwanted data can enhance the accuracy of the system. But the processing of missing values will also lead to a negative impact on results since there is a high risk of losing valuable information. Figure 1 shows the overall process of thyroid disease prediction with modified QPSO and QSVM.

Further, the feature selection method is implemented to identify the important features which are relevant for the prediction and classification of the output. The modified QPSO is employed to increase the exploitation and exploration properties for function optimisation. It also enhances the global and local search capabilities of evolutionary methods and involves tuning of parameters. After the feature selection method, the data is embedded into the train and test split. Then the QSVM is deployed in the training dataset for better classification of thyroid-affected persons from a normal person.

Figure 2 show the Modified QPSO and QSVM. This approach tends to improve the training speed by increasing the computational power. They convert the classical data into quantum states and build the kernel of SVM based on the quantum states. The performance of QSVM is compared with RF, SVM and KNN to evaluate the efficiency of the proposed model. The classified data is then fed to the prediction phase, which compares the trained model with the test data, and the output is predicted. The estimated output is evaluated based on derivative parameters, such as F1-score, recall, precision and accuracy.

3.1 | Modified QPSO (quantum particle swarm optimisation)

The QPSO is considered a supervised data clustering method. Learning accuracy, over-fitting, increased model learning and evaluation time are all performed in the feature selection method. It is defined as the method of choosing the significant feature from a dataset and eliminating unnecessary or redundant data to improve classification accuracy by decreasing the processing cost. In the traditional PSO, individuals in the population are known as particles and tend to move in search space, and each individual is a candidate solution. The movement of particles is the composition of two random weight and velocity influences to move towards the best position. But they lack in producing better quality solutions; producing early convergence and updating of velocity in memory are necessary. So a modified QPSO is implemented in the proposed system for better feature selection. Figure 3 shows the overall process involved in modified QPSO. The modifications added in the QPSO are adaptive balance with alpha, balance among the intensification and differentiation of gamma and re-initialization of poor performing population.

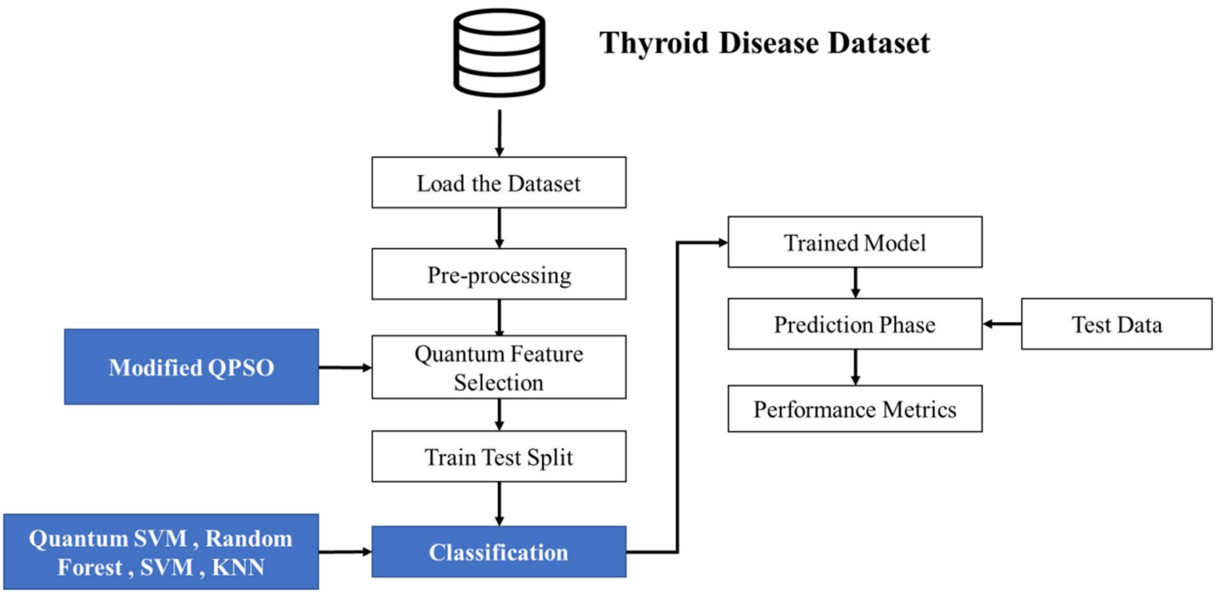

FIGURE 1 Overall process of thyroid disease prediction with modified QPSO and QSVM.

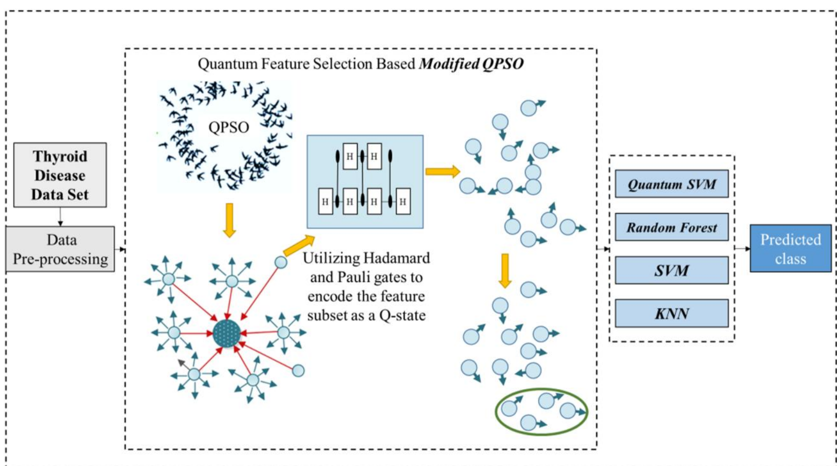

FIGURE 2 Modified QPSO and QSVM.

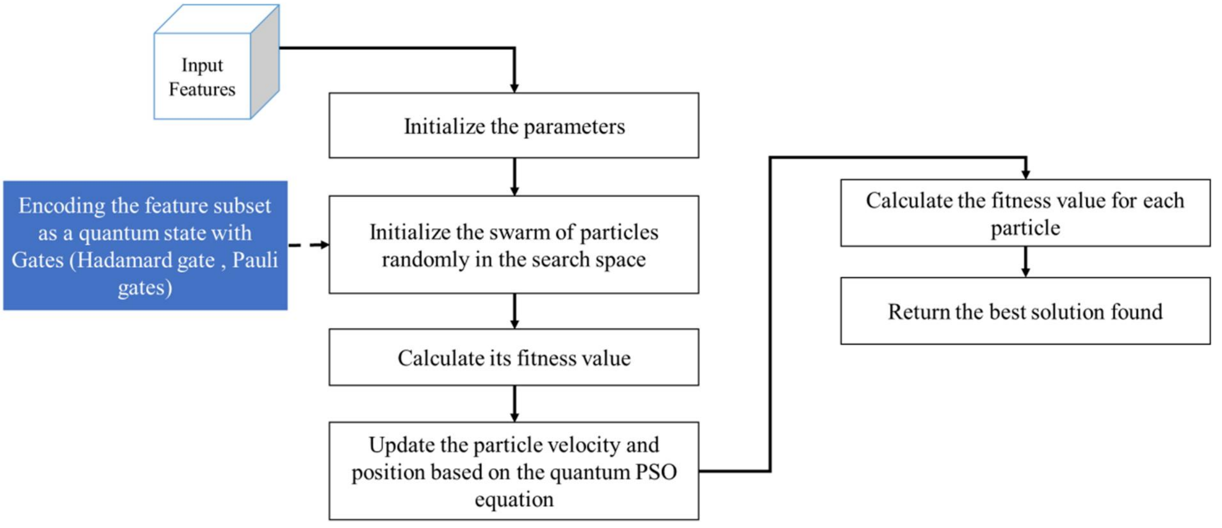

FIGURE 3 Feature selection with modified QPSO.

Figure 3 shows the feature selection with modified QPSO. The input from the loaded dataset is fed for the feature selection method and the parameters are initialised. Then the search space of random swarm particles is selected by using the gates such as Hadamard and Pauli gates. These gates are used to encode the features into quantum states for selecting enhanced optimal search space. Further, the fitness value is calculated. The solution quality in terms of classification accuracy is calculated by classifying the training datasets using selected quantum features. Feature cost and classification accuracy are the two important key factors used to evaluate the fitness function. The individual with high classification accuracy and a low total feature cost produces a high fitness value. This individual with high fitness value has a high probability of being selected for the next generation. Thus from the fitness value calculated, the particle velocity and position are fixed based on QPSO. Then the best optimal solution for feature extraction is selected. Therefore, the proposed modified QPSO tends to overcome the problem of poor exploration and getting stuck into local minima. It improves the exploitation and exploration attributes for function optimisation. A balance between global best and personal best positions based on the alpha parameter is employed, and the balance between diversification with the gamma parameter for intensification is performed. The $\propto$ and β represents the Hadamard gate and Pauli gates, respectively. This method is achieved by applying the following techniques.

3.1.1 | Hadamard gates

The respective system uses the parameter called Hadamard gates for balancing among the global best and personal position. This process makes the algorithm to be prominent on either global best or personal position, depending on the classification. Therefore, by utilising the alpha parameter, the proposed method manages both the global search and local search. Correspondingly, in the process of adaptive balance with alpha, the balance among the global best and personal best is processed using the alpha parameter. It personally guides the movement of particle for the best solution where the global best guide for the global best position.

The Hadamard gate is also defined as the H gate, which is used to alter, the qubit from a clustering state to a uniform superposed state. It is also expressed as 90° rotational gates (π), in the Bloch sphere about the axis. The Hadamard gate matrixes are symmetric and real, which is easy and simple to apply.

$$\propto = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \simeq R_{\vec{n}}(\pi),\ \vec{n} = \frac{1}{\sqrt{2}}\,(1,0,1) \tag{1}$$

3.1.2 | Pauli gates

Pauli gates are considered useful and are slightly delicate since the non-integer powers are not unique. It consists of a diagonalisable matrix with two square root values of any numbers along eigenvalues having 2^n square roots. This is defined by using the Pauli rotation gates, and π is denoted as the rotation in the Pauli gate up to phase and is given by Equation (2),

$$R_x(\pi) = e^{-k\,\frac{\pi}{2}x} = -kx \tag{2}$$

The powers of Pauli matrices are denoted by the following equation:

$$\beta = e^{-k\,\frac{\pi}{2}t(x-1)} \simeq R_x(\pi t) \tag{3}$$

Based on Equation (4), it is represented that the Pauli-power gates, the spin states in a similar way nearby Bloch sphere similar to Pauli-rotation gates. These Pauli matrices also possess some computational advantages.

By implementing infinite iterations, the modified QPSO algorithm can achieve an optimal solution. To alleviate the issue of particles, caught in a local position, re-initialisation of a part of the population produces new outcomes and aids the particles to get escaped from the local position and accelerate towards a global best solution.

Here, the velocity v_k, position x_k, and population of particles with dimension dim of each particle at iteration $(1 \leq \dim \leq D)$ at $(t+1)^{th}$ is given by the following equations:

$$x_k\ (t) = [\ x_{k,1}(t), x_{k,2}(t), \ldots, x_{k,\dim}(t)\ ] \tag{4}$$

$$v_k\ (t) = [\ v_{k,1}(t), V_{k,2}(t), \ldots, v_{k,\dim}(t)\ ] \tag{5}$$

The enhancing of position and velocity of kth' particle for dimension $k(1 \leq k \leq N)$ at iteration $(t+1)$ is represented by the following equations:

$$\begin{aligned} v_{k,\dim}\ (t+1) = {} & \mathrm{weg}.v_{k,\dim}(t) \\ & + c_1.r_1.\left(\mathrm{pbest}_{k,\dim}(t) - x_{k,\dim}(t)\right) \\ & + c_2.r_2.\left(\mathrm{gbest}_{\dim}\ (t) - x_{k,\dim}\ (t)\right) \end{aligned} \tag{6}$$

$$x_{k,\dim}(t+1) = x_{k,\dim}(t) + v_{k,\dim}(t+1) \tag{7}$$

Where weg denotes the inertia weight, r1 and r2 ϵ (0, 1), c1, c2 are constants that are dispersed continuously. The proportion of poor performing population is re-initialized to support escape local optima. This local optima are the poor solution. Hence, in order to stop the system from falling into the local optima, the amount of poor performing operations is re-initialized. From that, the respective model presents new search points which directs to the finest solutions. This will enhance the search for global optima. The particle global best position is given by *gbest* and the best position is given by $pbest_i$. Each particle converges into local attractor p_k and is represented by the following equation:

$$p_k(t) = \left[p_{k,1}(t), p_{k,2}(t), \ldots p_{k,\dim}(t)\right] \tag{8}$$

$$\text{Where } p_{k,\dim}(t) = \left(\varnothing_{1,\dim} * pbest_{k,\dim}(t) + \varnothing_{2,\dim} * gbest_d(t)\right) / \left(\varnothing_{1,\dim} + \varnothing_{2,\dim}\right) \tag{9}$$

and $\varnothing_{1,\dim}$ and $\varnothing_{2,\dim}$ ϵ rand $(0,1)$ are dispersed consistently.

Here, quantum delta potential is employed to limit the drive of particles toward the attractor particle with p_k. These particles produce quantum behaviour by replacing Hilbert space, where the particle is denoted as the wave function, which hinges on particle position. Thus the updated position is denoted by the following equation:

$$x_{k,\dim}(t+1) = p_{k,\dim}(t) \pm \frac{L_{k,\dim}(t)}{2} \ln\left(\frac{1}{u_{k,\dim(t)}}\right) \tag{10}$$

Where $u_{k,\dim}$ ϵ r *and* (0, 1) and is dispersed consistently and the delta potential length $L_{k,}$ is denoted by the following equation:

$$L_{k,\dim}(t) = 2\,\beta.\ |p_{k,\dim}(t) - x_{k,\dim}(t)| \tag{11}$$

In which β is defined as the Pauli matrix parameter regulating the speed of convergence speed. The average position of swarm best position is given by *mbest* position and is denoted by the following equation:

$$\begin{aligned} mbest\ (t) = {} & \frac{1}{N}\sum_{k=1}^{N} pbest_{k,1}(t), \frac{1}{N}\sum_{k=1}^{N} pbest_{k,2}(t), \ldots, \\ & \times \frac{1}{N}\sum_{k=1}^{N} pbest_{k,\dim}(t) \end{aligned} \tag{12}$$

$$\begin{aligned} x_{i,d}(t+1) = {} & p_{i,\dim}(t)\ \pm\ \beta.\ \left|mbest_{\dim}(t) - x_{k,\dim}(t)\right| \\ & .\ln\left(\frac{1}{u_{i,\dim}(t)}\right) \end{aligned} \tag{13}$$

The mean of best position helps optimise the collaboration of particles with outcomes in the enhanced global search capability. From the algorithm, it is found that the grouping of particles restricts the search pattern, which leads to sub-optimal solutions. The random weights are fed to global best particles and personal best particles. Here, the p_k is evaluated by attributing more weights to the best particle $pbest_i$ in the first iterations, and more weight is attributed to *gbest*. The local attractor particle $\propto$ is given by the following equation:

$$p_k = \propto * r_1 * pbest_k(t) + (1 - \propto) * r_2 * gbest(t) \tag{14}$$

Where r_1 and $r_2 \epsilon$ r *and* (0, 1), which starts the movement around $pbest_i$ and *gbest*. The value of $\propto$ decreases and is given by the following equation:

$$\propto = \mathrm{weg}_1 + \left(\mathrm{weg}_2 - \mathrm{weg}_1\right) * \left(\frac{t}{\mathrm{Max}_{\mathrm{iter}}}\right) \tag{15}$$

The other variable γ is given by the following equation:

$$\gamma = \beta * \log\left(\frac{t}{u_k(t)}\right) \tag{16}$$

The gamma parameter balances the intensification and differentiation in the system. Here, the differentiation signifies the searching of an extensive kind of search space to identify a potential solution where the intensification is centred on exploiting regions in the search space and enhancing the solution. Therefore, by using the gamma, the proposed model can manage the balance among the intensification and differentiation. A greater value of the gamma upholds differentiation. It enhances the exploration of search space. Contrarily, the lesser value of gamma upholds the intensification by directing the system to focus on exploiting the important region for the purpose of improving the solution.

Where β is the time fluctuation and reduces the linearity from 2 to 0, it is spread consistently. To measure the adequate jump away from the present position, the X_{rand} of the population is selected with length L calculated. It is the absolute distance of the present particle from the selected random particle.

$$L = \gamma * \text{abs}(C * X_{\text{rand}}(t) - x_k(t)) \tag{17}$$

$$x_k(t+1) = x_{\text{rand}}(t) \pm \gamma * \text{abs}(C * x_{\text{rand}}(t) - x_k(t)) \tag{18}$$

Equation (18) includes C, which is a random variable. It permits adjustable step size to carefully display the search space. During the process of exploitation, particle meets together upon global minima position.

Based on the value gamma (γ), the decision to exploit and explore ranges from 0 to 1. In case if the value of gamma is higher than the pre-defined threshold, the particles explore or perform exploitation. The updated positions of the present particle are given by Equations (20) and (21).

$$L = \gamma * \text{abs}\left(C * p_k(t) - x_k(t))\right) \tag{19}$$

$$x_k(t+1) = p_k(t) \pm \gamma * \text{abs}\left(C * p_k(t) - x_k(t)\right) \tag{20}$$

The gamma plays a significant role in final convergence and ensuring exploitation. These three modifications $\propto$, β, γ together play a decisive role to optimise the performance of the modified QPSO approach.

Correspondingly, adaptive balance among the global and personal finest solution through alpha, balance of intensification and differentiation through gamma and re-initialization of poor performing population improves the performance of the proposed system. This changes achieve better balance of exploitation and exploration, enhancing the escape technique and adaptability in the respective system. Therefore, utilising modified PSO and quantum operators in the proposed research feature selection enhances the capability of using higher feature space. Further, identifying the combination of feature directs to the enhanced performance of the proposed method.

3.2 | QSVM (quantum Support Vector Machine)

QC enables computation advantages which allows classification in higher dimensions. The feature selected from QPSO is fed as input to QSVM for classification. The QSVM is also trained similarly to the classical method only when the kernel is obtained. The proposed method used RBF (Radial Basis Function) as the kernel function for the QSVM system. The major advantages of using RBF are high-dimensional feature space, speed and reduced overfitting of data. Accordingly, the RBF kernel maps the input data into the feature space of high dimension to identify complex relationship in the data. Besides, quantum computing further improves the capacity by a greater dimensional mapping for the identification of complex patterns. In addition, it uses non-linearity in the decision boundaries and effective calculation of the complex values of the kernel supports examination of non-linear relationships. Further, the attributes of quantum enhance the speed in the calculation and evaluation of the kernel. Moreover, the efficiency of quantum computing in the transformation of data improves the capability of generalisation. Furthermore, the quantum computing produces effective quantum feature maps which improves the data transmission and enhances the performance of the RBF kernel.

The kernel function-based QSVM is a typical learning algorithm in which the SVM tends to increase the learning ability by determining the kernel function. The selection of different kernels produces different QSVM, along with different kernel parameter functions. The QSVM classifying model is trained by taking the feature map along with the training and testing set for each fold and applying the quantum kernel, which allows classification with less computational cost. It tends to reduce complexity when compared with other models. The use of QC enhances the evaluation of the inverse matrix. Thus the QSVM model decreases the loss by optimising its parameters. The steps involved in QSVM are shown in Algorithm 1.

Algorithm 1 QSVM

```
Input: Thyroid disease dataset
Output: Predicted class label
1. Data Encoding:
   Encode features from QPSOinto quantum
   states
   using quantum encoding.
2. Quantum Kernel Computation:
   Compute the quantum kernel matrix K using
   the encoded quantum states.
   Define the hyperparameters of the quantum
   kernel function.
   Quantum kernel type: Gaussian
   Quantum kernel parameters: sigma for
   Gaussian and
   degree for polynomial.
3.Quantum Support Vector Classification with
RBF Kernel:
   Define the hyperparameters C (penalty
   parameter) and γ (gamma) for the SVM.
   Initialize the weight vector w and bias b.
   Initialize the quantum state for the
   support vector register.
```

```
For each training data:
  Initialize the quantum kernel value QK
  as 0.
  Calculate the quantum kernel value
  For each support vector quantum state:
    Compute the quantum kernel value
    QKibetween the ith training data and
    the support vector
    quantum state using the fixed RBF kernel
    formula with γ.
    QK += QKi
  Calculate the predicted class label
  PredictedClass = Sign(w · QK + b)
  Calculate the true class label
  TrueClass = True class label for the
  ithtraining data
  Calculate the error term
  ErrorTerm = TrueClass - PredictedClass
  Update the weight vector w and bias b
  using the error term and quantum kernel
  value
  w += C * ErrorTerm * QK
  b += C * ErrorTerm
4. Quantum Measurement and Prediction:
  Apply a quantum measurement to collapse
  the support vector quantum
  state and obtain the predicted class
  labels for X_test.
  Assign the predicted class labels.
```

The proposed modified QPSO system is applied for enhanced feature selection by selecting the global best position of particles by selecting the relevant data and eliminating redundant data. With the use of QSVM, the classification accuracy of thyroid diseases is optimised by using the optimal selected features. Compared with traditional approaches, the proposed method produces improved advantages.

4 | RESULTS AND DISCUSSION

The results produced by the implementation of the proposed system are discussed in this section. Additionally, dataset description, EDA, performance analysis, identified outcomes and comparative analysis of conventional methods are presented.

4.1 | Dataset description

The study has considered a thyroid disease dataset consisting of 29 attributes. Nearly 2800 training data and 972 test data are used. The attributes of thyroid disease dataset is represented in Table 1.

The proposed method used the attributes of 3, 4, 6, 8, 10, 11, 14, 17, 18, 19, 20, 21, 23, 26, 27 and 28. Correspondingly, gender of the patient is used in the model as it can impact the clinical representation of the thyroid disease because the levels of hormone and the symptoms differ based on gender. Similarly, the medication of thyroxine is an essential feature in the management of thyroid disease. In the same way, the level of T3 is a significant measure to differentiate the type of thyroid disease such as hypothyroidism or hyperthyroidism. Likewise, level of FTI plays a major role in signifying the value of thyroxine in the blood. In the same way, pregnancy related hormonal changes can affect the function of thyroid. Correspondingly, TBH levels plays a major role in the availability of thyroid hormone. Moreover, TSH value is used to detect the existence of feature whether the TBG, T3 and TSH levels highlights the importance of thyroid disease classification.

4.2 | Performance metrics

To estimate the performance of the proposed system, three major parameters are used: precision, accuracy, F1 score, and recall.

A. Precision: The precision of the system is represented as the capability of the classifier to correctly predict the positive class and is given by the Equation (21),

$$\text{Precision} = \frac{\text{Tp}}{\text{Tp} + \text{Fp}} \tag{21}$$

Where Tp denotes True positive and Fp is False positive.

B. Accuracy: The accuracy of the system is defined as the ratio of the total number of correctly classified models to the total number of classified models and given by Equation (22)

$$\text{Accuracy} = \frac{\text{Tp} + \text{Tn}}{\text{Tp} + \text{Fp} + \text{Fn} + \text{Tn}} \tag{22}$$

Where Tp represents True positive, Tn is True negative, Fp is False positive and Fn is False negative.

C. Recall: Dividing the number of times the classifier identified the negative class by all the times the negative class is obtained is expressed in Equation (23),

$$\text{Recall} = \frac{\text{Tp}}{\text{Tp} + \text{Fn}} \tag{23}$$

D. F1 Score: It is a metric that depends on precision and recall and is given by Equation (24),

$$\text{F1 Score} = 2 + \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{24}$$

E. ROC Curve: To visualise the efficacy of the respective system for binary classification, ROC is used. It is the representation of producing continuous variables of

TABLE 1 Attributes in thyroid disease dataset.

1	Class	The objective characteristic: Euthyroid, hyperthyroid or hypothyroid.
2	Age	Age in years for patients
3	Sex	Patients age (male or female).
4	Thyroxine	Taking the medicine of thyroxine (Yes or No)
5	Level of TSH	The level value of TSH
6	Level of T3	The level value of T3
7	Level of T4	The level value of T4
8	FTI	T4: Free index.
9	Level of TBG	The level of thyroxine-binding globulin
10	TSH level_calculated	TSH level calculated: (Yes or No).
11	T3 level_ calculated	T3 level calculated: (Yes or No).
12	T4 level_ calculated	T4 level calculated: (Yes or No).
13	FTI level_ calculated	FTI level calculated: (Yes or No).
14	TBG level_ calculated	TBG calculated: (Yes or No).
15	Source-referral	Patient's referral source (GP, endocrinologist, other).
16	Hypothyroidism-query	Patient mentioned on expected hypothyroidism (Yes or No).
17	Hyperthyroidism-query	Patient mentioned on expected hyperthyroidism (Yes or No).
18	State of sick	State of patient presently ill or not (Yes or No).
19	State of pregnant	Patient pregnant or not (Yes or No).
20	Hypothyroid-TSH	Level of TSH in hypothyroidism for patients.
21	Hyperthyroid-TSH	Level of TSH in hyperthyroidism for patients.
22	Hypothyroid-T3	Level of T3 in hypothyroidism for patients.
23	Hyperthyroid-T3	Level of T3 in hyperthyroidism for patients.
24	Hypothyroid-T4	Level of T4 in hypothyroidism for patients.
25	Hyperthyroid-T4	Level of T4 in hyperthyroidism for patients.
26	Hypothyroid-FTI	Level of FTI in hypothyroidism for patients.
27	Hyperthyroid-FTI	Level of FTI in hyperthyroidism for patients.
28	Hypothyroid-TBG	Level of TBG in hypothyroidism for patients.
29	Hyperthyroid-TBG	Level of TBG in hyperthyroidism for patients.

specificity and sensitivity. The ROC curve is an evaluation metric in which the curve is plotted between TP (True Positive) and FP (False Positive) at various threshold levels. The performance is thus demonstrated by using a confusion matrix.

4.3 | Performance analysis

The results produced by the proposed model denote the behaviour and accuracy of the method implemented. Several attributes based on the thyroid disease dataset are implemented to produce better feature selection and classification accuracy. Table 2 denotes the parameters used for feature selection methods using QPSO.

The parameters used by QPSO for feature selection are N particle, maximum iterations, number of dimensions, bounds, g-acceleration co-efficient, number of qubits and selected features. The Table 3 represents the precision of different conventional algorithms and the proposed QSVM.

From Figure 4, it is inferred that the proposed QSVM produced better accuracy rates in terms of recall, F1 score, precision and accuracy. The accuracy is found to be 98.77, the precision is 99, the recall is 99 and F1-score is 99. Whereas, the accuracy rates of RF, KNN, and SVM are 50.17%, 92.74%, and 86.71%, respectively. Figures 11 and 12 represent the ROC and confusion matrix of the proposed QSVM.

From the Table 3 and Figure 4, internal comparative analysis uses the metrics such as recall, f1-score and precision because they are the critical calculation metrics for imbalanced process such as classification of thyroid disease. It provides insights into the capability of the system in accurately classifying the positive and negative cases. Besides, balancing among the false positive and true positive identification. The proposed system overcomes other conventional algorithms in terms of effective handling of complex patterns, non-linearity and better feature selection. Correspondingly, QSVM uses the quantum kernel function to identify complicated patterns. This kernel function converts the data into high dimensional feature space, which improves the system capability of handling complex patterns. Besides, the kernel function in QSVM outperforms the traditional algorithms in capturing non-linear similarity in the respective system. Further, combination of modified QPSO assists the QSVM to identify essential features. The centred feature subset improves the proposed model accuracy by minimising the noise and neglecting unwanted data.

Correspondingly, as an outcome of the internal comparison, the proposed QSVM system produced better results in the classification than the RF, KNN and SVM. The attributes of quantum in dimensionality minimisation, better computation of the kernel and the quantum feature mapping improves the performance of the proposed model than the RF, KNN and SVM models. Similarly, QSVM system uses quantum circuits and gates to perform complicated processes such as matrix inversions and conversions. It enhances the computation speed in the kernel operation. Besides, QSVM influences the quantum parallelism to continuously operate several data and calculate complicated conversion, and the quantum feature map allows effective representation in space of high dimension. It achieves better computation as the quantum computers process quantum parallelism to compute inner vector products in quantum state. This calculation enhances the computation in the system. In addition, the strength of effective computation enhances the dimensionality reduction in the system by recollecting significant data which improves the classification performance. Finally, quantum permits significant data processing and representation which contribute in attaining faster outcome when compared to RF, KNN and SVM.

The proposed system uses QSVM on the dataset for thyroid disease classification. The interpretation of traditional methods with the proposed method is performed, and the results obtained are shown in the form of ROC and confusion matrix. The ROC gives the probability curve at various threshold values and specifically separates the signal from noise. It also predicts the performance of the model at all classification thresholds. Whereas, the confusion matrix is the summarised table used to assess the performance of the classification module. Figures 4 and 5 denote the ROC and confusion matrix of the RF classification model.

The RF classifier has been implemented, and the outcomes are represented in the form of ROC and confusion matrix. From Figures 5 and 6, it is denoted that, the ROC value needs further enhancement as poor performance might lead to overfitting issues, which deviate the model from producing appropriate accuracy. Further, confusion matrix rate also needs further improvement as this classifier is not able to distinguish between positive and negative values. It predicts a constant or random class for every data point. The confusion matrix denotes the major misclassification of positive and negative labels. Figures 7 and 8 denotes the ROC and confusion matrix of the K-NN classifier.

TABLE 2 Parameters used for feature selection methods using QPSO.

QPSO parameters	
N Particle	40
MaxIters	1000
No. Dim	30
Bounds	−2.56, 5.12
g- acceleration co-efficient	0.96
num_qubits	30
Selected features:	3, 4, 6, 8, 10, 11, 14, 17, 18, 19, 20, 21, 23, 26, 27, 28

TABLE 3 Performance of traditional models and proposed QSVM.

MODEL	Accuracy	Precision	Recall	F1-score
Random forest	50.17	63	50	35
KNN	92.74	93	93	93
SVM	86.71	87	87	87
QSVM	98.77	99	99	99

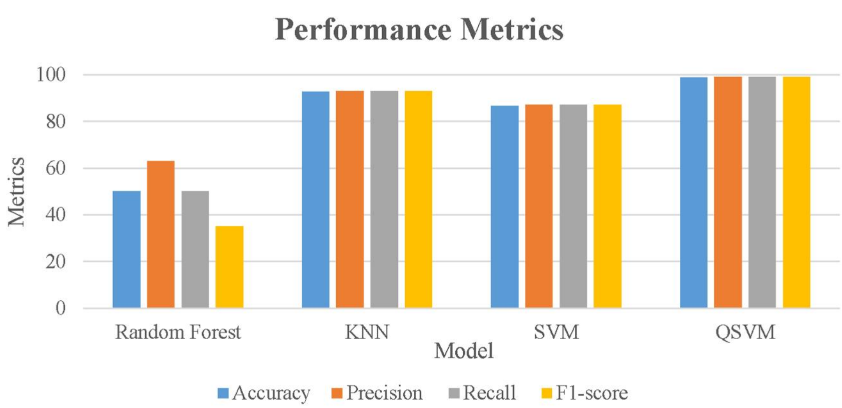

FIGURE 4 Graphical representation of traditional methods with QSVM.

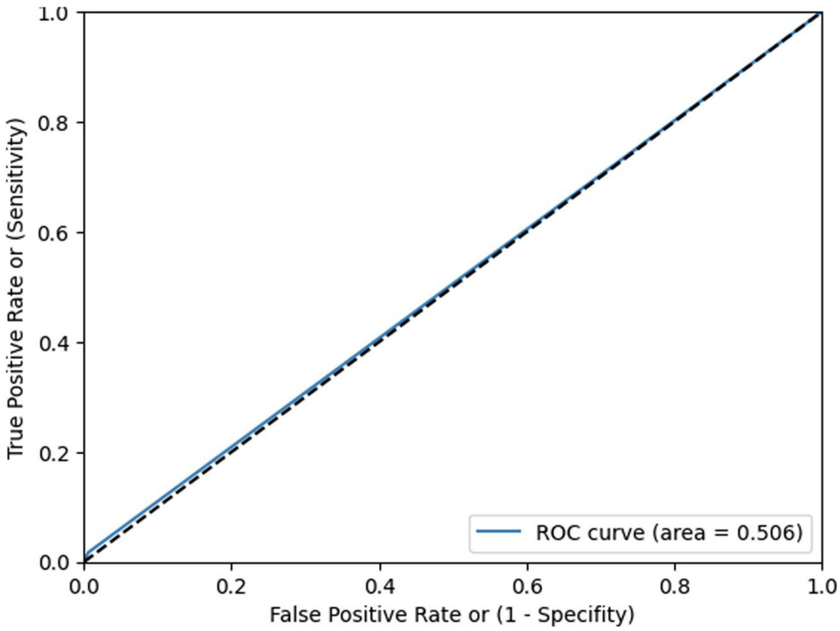

FIGURE 5 ROC for random forest model.

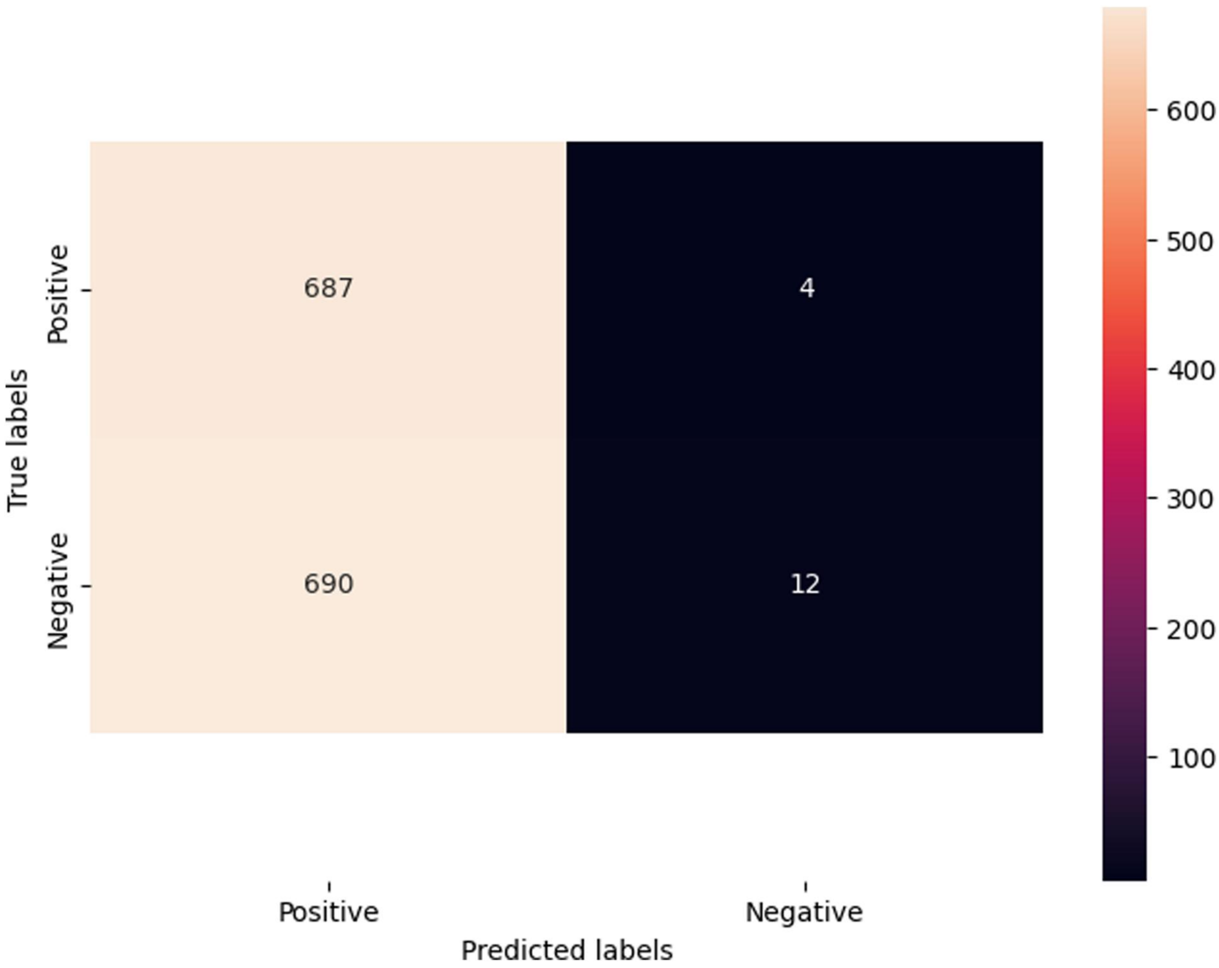

FIGURE 6 Confusion matrix for random forest.

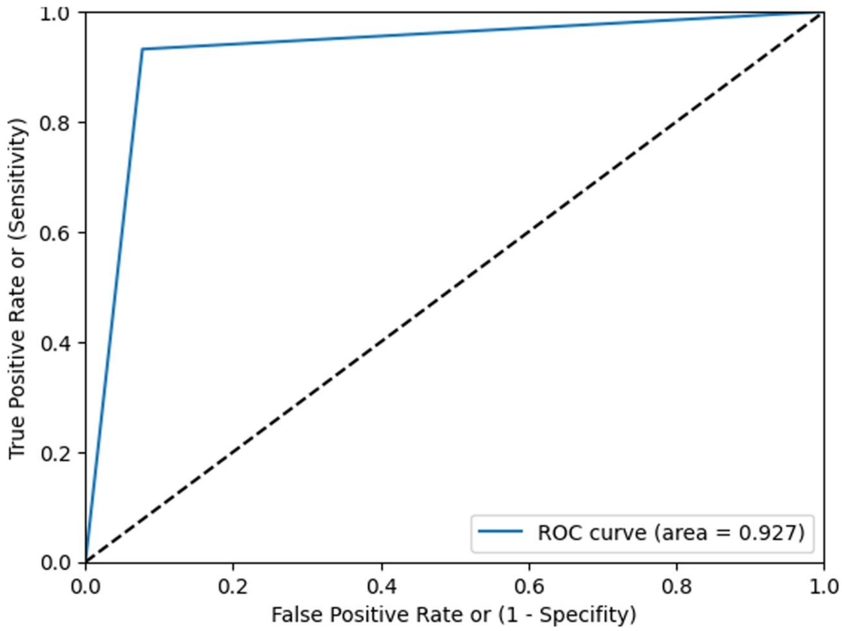

FIGURE 7 ROC of K-Nearest neighbour Model.

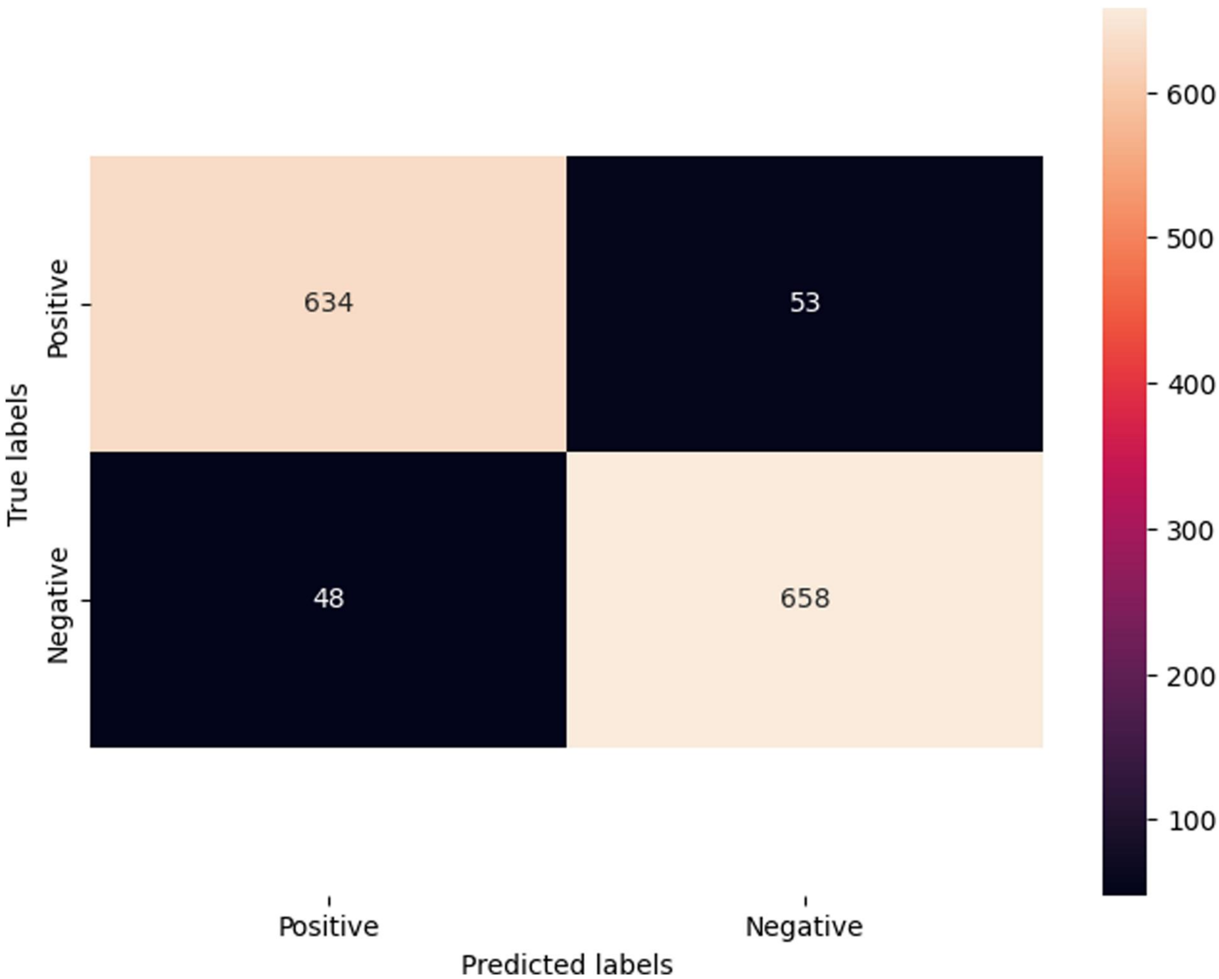

FIGURE 8 Confusion matrix for K-Nearest Neighbour model.

The ROC of the KNN model denotes that there is an improved chance that the classifier can discriminate more between a true positive and true negative than a false positive and false negative. Then, improper classification of positive and negative labels is found in the confusion matrix of the K-NN model. It has misclassified nearly 53 and 48 data which lead to the wrong prediction of thyroid disease. Figures 9 and 10 show the ROC and confusion matrix of SVM.

From the figure, it is found that the SVM model possessed normal disease prediction, and there is a chance of misclassification. The confusion matrix also denotes that nearly 100 and 85 data are wrongly predicted. This wrong prediction of thyroid disease will lead to unnecessary involvement in treatment, which tends to change the lifestyle of a normal non-thyroid-affected person.

From Figures 11 and 12, it is inferred that the proposed system produces greater ROC and minimum error prediction, which intimates that the proposed QSVM classification method performs better than the other conventional methods. The classifier tends to correctly distinguish between all positive and negative class points. Thus, it is denoted that the traditional approaches of feature selection and classification models lack accuracy and misclassification of thyroid diseases. Only a few studies applied QC in feature selection and classification methods based on binary classification. So, to overcome this, the proposed method is implemented. Here, the modified QPSO is employed instead of other methods, as it tends to improve in selecting the global best particle position to produce a better feature selection algorithm. And QSVM is used instead of traditional SVM, as it increases the training speed of larger datasets with a better generalisation approach. It also possesses the ability to handle non-linear data as they use quantum kernel functions, which do not apply to traditional SVM methods. Therefore, the modified QPSO and QSVM produce better accuracy when compared with other conventional methods.

5 | DISCUSSIONS

Numerous existing research studies focused on the classification of thyroid nodules, thyroid ailments and thyroid lesions. Correspondingly, limited studies centred on thyroid disease. For instance, the existing method used ML based algorithms for the classification of thyroid ailments through clinical dataset [41]. Similarly, traditional model used multi-objectives optimisation of features selection for the classification of thyroid nodules [12]. In the same way, the ML-based system was used for the classification of thyroid types such as hypothyroid and hyperthyroid [22]. In the classification of thyroid disease, several conventional methods focused on clinical dataset and medical dataset. Moreover, dataset used by several conventional classification methods are the smaller datasets. Conversely, the proposed model used thyroid disease dataset for the classification of thyroid because it is the largest dataset and it is focused on binary classification of thyroid and non-thyroid. However, dataset utilised by the respective model is insufficient in several existing models. Due to the absence of suitable datasets in the traditional research studies, external comparison is not feasible, which would impact the reliability of comparison. Therefore, the internal comparison is executed in the proposed method through Random Forest, KNN and SVM [42].

6 | CONCLUSION

Thyroid is one of the most critical diseases and has the potential to be transformed into a common disease. So, early prediction of the disease helps people to overcome this issue. There are various approaches implemented in the analysis of thyroid disease prediction based on classification approaches. Nevertheless, these techniques suggested decreasing the

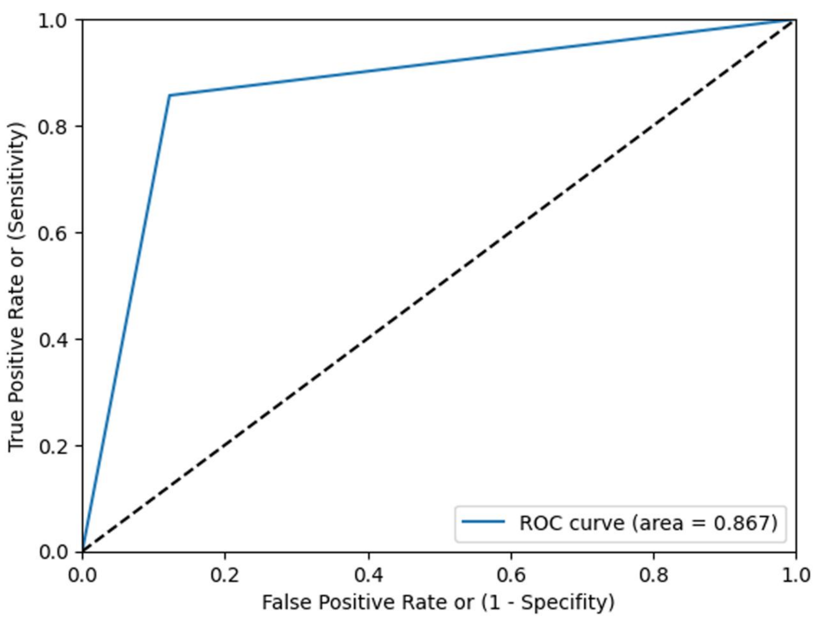

FIGURE 9 ROC for support vector machine.

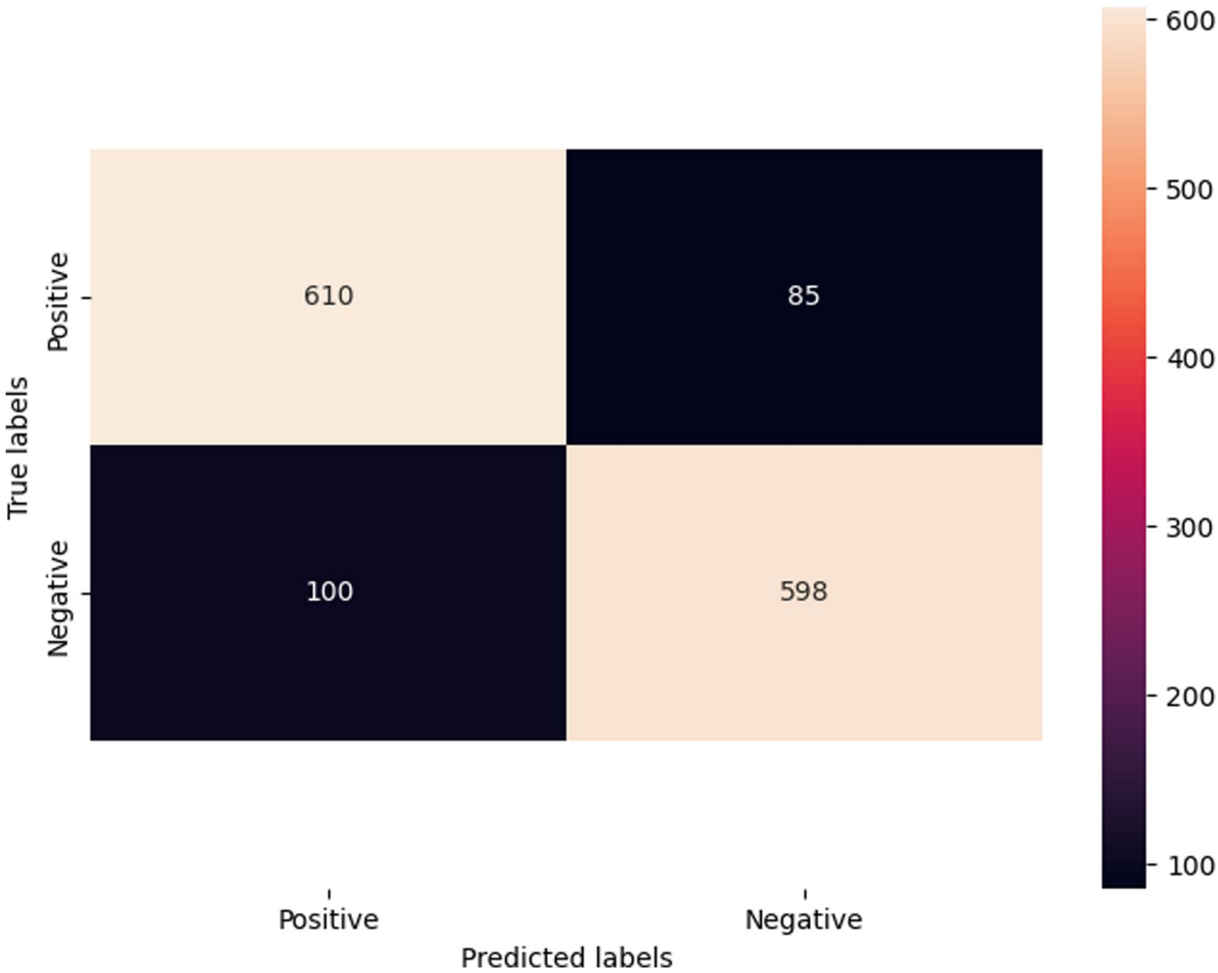

FIGURE 10 Confusion matrix for support vector machine.

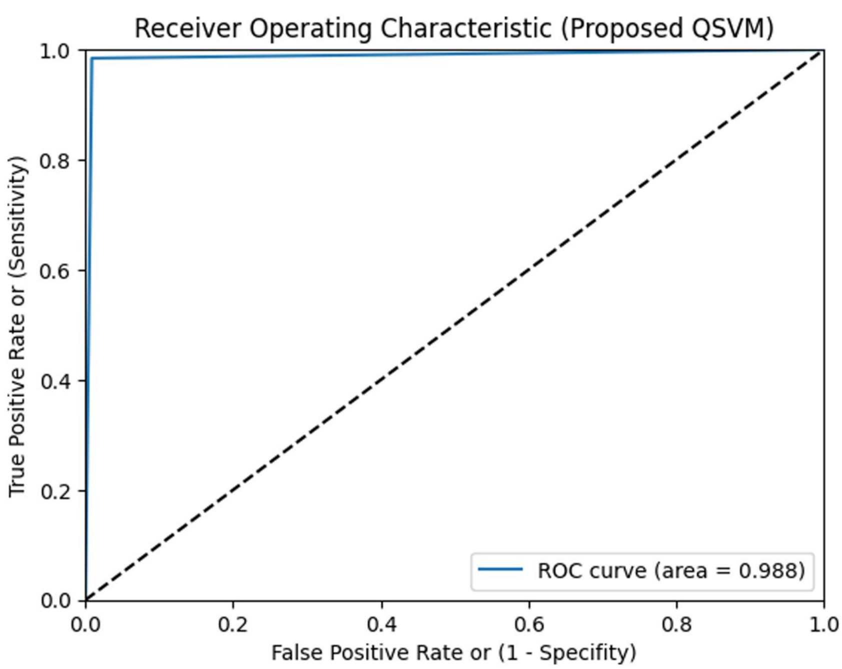

FIGURE 11 ROC for proposed quantum support vector machine.

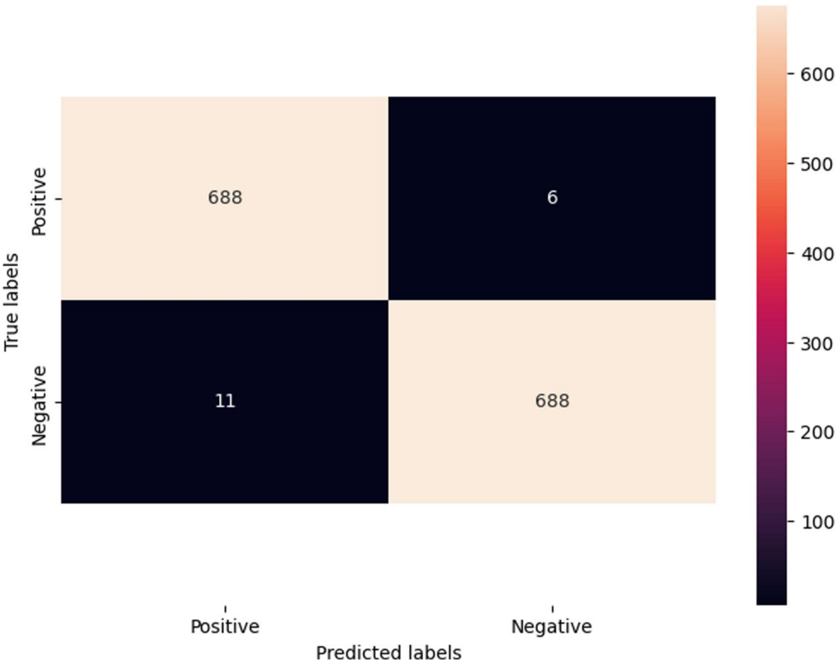

FIGURE 12 Confusion matrix for proposed quantum support vector machine.

number of parameters used by patients for diagnosis of thyroid disease. Hence both QC and ML algorithms were combined to enhance the prediction and classification of thyroid-affected persons and non-thyroid-affected persons. The proposed model used QPSO for the feature selection as it was the effective metaheuristic algorithm which effectively functioned to attain an optimal feature subset through vital information in an optimal time. However, it can easily fall into local optimum in high dimensional space. To resolve the limitation of traditional PSO and to enhance the performance, the proposed system used modified QPSO to enhance the search optimisation of particle position from local optimum to the global best position.

Further, the ML-based classification method was implemented to optimise the speed of computational power and accuracy of the classifier. But the conventional SVM produced overlapping while using large datasets. Further, the proposed method was associated with supplementary conventional procedures to evaluate the performance of the model. From the comparison, it was found that the proposed model produced 98.77% accuracy, which was better than the other models. This improved accuracy prediction will help doctors to predict thyroid disease in the primary stage, which benefits people to maintain a healthy society. Then, the efficacy of the system was measured based on derivative attributes, such as recall, F1 score, accuracy and precision. In the future, the proposed system can be enhanced by applying several feature selection and classification algorithms to improve the performance of thyroid disease prediction with better accuracy rates.

The proposed method is intended to contribute to the medical field by assisting qualified doctors with improved accuracy in the detection of thyroid disease. Likewise, it captures complex patterns and similarity within the data of thyroid disease. In the same way, the respective model is expected to enhance the thyroid disease treatment through primary detection mechanism. Further, the proposed model is planned to contribute in the clinical practice through effective decision-making tool for the qualified physicians. To conclude, the proposed system is envisioned to contribute in the medical industry for thyroid diagnosis and to increase the survival rate of the thyroid disease.

AUTHOR CONTRIBUTIONS

Mohemmed Sha: Conceptualization; Investigation; Methodology; Validation; Writing – original draft; Writing – review & editing.

ACKNOWLEDGEMENTS

This study is supported via funding from Prince sattam bin Abdulaziz University project number(PSAU/2023/R/1445).

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Mohemmed Sha https://orcid.org/0000-0002-8181-4887

REFERENCES

1. Rebentrost, P., Mohseni, M., Lloyd, S.J.: Quantum support vector machine for big data classification. Phys. Rev. Lett. 113(13), 130503 (2014). https://doi.org/10.1103/physrevlett.113.130503
2. Zhou, M.-G., et al.: Experimental quantum advantage with quantum coupon collector (2022)
3. Zhou, N.-R., et al.: Hybrid quantum–classical generative adversarial networks for image generation via learning discrete distribution. Signal

Process. Image Commun. 110, 116891 (2023). https://doi.org/10.1016/j.image.2022.116891
4. Bai, Q., Hu, X.J.Q.I.P.: Quantity study on a novel quantum neural network with alternately controlled gates for binary image classification. Quant. Inf. Process. 22(5), 184 (2023). https://doi.org/10.1007/s11128-023-03929-y
5. Wu, C., et al.: Quantum SUSAN edge detection based on double chains quantum genetic algorithm. Phys. Stat. Mech. Appl. 605, 128017 (2022). https://doi.org/10.1016/j.physa.2022.128017
6. Huang, S.-Y., et al.: Image classification and adversarial robustness analysis based on hybrid quantum–classical convolutional neural network. Opt Commun. 533, 129287 (2023). https://doi.org/10.1016/j.optcom.2023.129287
7. Gong, L.-H., et al.: Born machine model based on matrix product state quantum circuit. Phys. Stat. Mech. Appl. 593, 126907 (2022). https://doi.org/10.1016/j.physa.2022.126907
8. Zhou, N.-R., et al.: Quantum particle swarm optimization algorithm with the truncated mean stabilization strategy. Quant. Inf. Process. 21(2), 42 (2022). https://doi.org/10.1007/s11128-021-03380-x
9. Zeng, Q.-W., et al.: Conditional quantum circuit Born machine based on a hybrid quantum–classical framework. Phys. Stat. Mech. Appl. 618, 128693 (2023). https://doi.org/10.1016/j.physa.2023.128693
10. Zhou, N.-R., et al.: Quantum K-nearest-neighbor image classification algorithm based on KL transform. Int. J. Theor. Phys. 60(3), 1209–1224 (2021). https://doi.org/10.1007/s10773-021-04747-7
11. Shivastuti, H.K., Manhas, J., Sharma, V.: Performance evaluation of svm and random forest for the diagnosis of thyroid disorder. Int. J. Res. Appl. Sci. Eng. Technol. 9, 945–947 (2021)
12. Aboudi, N., Guetari, R., Khlifa, N.: Multi-objectives optimisation of features selection for the classification of thyroid nodules in ultrasound images. IET Image Process. 14(9), 1901–1908 (2020). https://doi.org/10.1049/iet-ipr.2019.1540
13. Chaganti, R., et al.: Thyroid disease prediction using selective features and machine learning techniques. Cancers 14(16), 3914 (2022). https://doi.org/10.3390/cancers14163914
14. Tharwat, A., Hassanien, A.E.: Quantum-behaved particle swarm optimization for parameter optimization of support vector machine. J. Classif. 36(3), 576–598 (2019). https://doi.org/10.1007/s00357-018-9299-1
15. Rasheeduddin, S., Kumari, C.V.: A Noval System for Early Detection of Thyroid with Graph Cluster ant Colony Optimization
16. Shi, Y., et al.: Prediction of progression in idiopathic pulmonary fibrosis using CT scans at baseline: a quantum particle swarm optimization-Random forest approach. Artif. Intell. Med. 100, 101709 (2019). https://doi.org/10.1016/j.artmed.2019.101709
17. Chaganti, R., Rustam, F., De La Torre Díez, I., Mazón, J., Rodríguez, C., & Ashraf, I.: Thyroid disease prediction using selective features and machine learning techniques. Cancers 14(16), 3914. (2022) https://doi.org/10.3390/cancers14163914, In: s Note: MDPI stays neutral with regard to jurisdictional claims in published …
18. Yadav, D.C., Pal, S.: Discovery of hidden pattern in thyroid disease by machine learning algorithms. Indian J. Public Health Res. & Develop. 11(1), 61–66 (2020). https://doi.org/10.37506/v11/i1/2020/ijphrd/193785
19. More, K.: Classification of Thyroid Disease Using Machine Learning (2021)
20. Xi, N.M., Wang, L., Yang, C.: Improving the diagnosis of thyroid cancer by machine learning and clinical data. Sci. Rep. 12(1), 11143 (2022). https://doi.org/10.1038/s41598-022-15342-z
21. Alyas, T., et al.: Empirical method for thyroid disease classification using a machine learning approach. BioMed Res. Int. 2022, 1–10 (2022). https://doi.org/10.1155/2022/9809932
22. Umar Sidiq, D., Aaqib, S.M., Khan, R.A.: Diagnosis of various thyroid ailments using data mining classification techniques. Int. J. Sci. Res. Coput Sci. Inf. Technol. 5, 131–136 (2019). https://doi.org/10.32628/cseit195119
23. Guleria, K., et al.: Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning. Measurement: Sensors 24, 100482 (2022). https://doi.org/10.1016/j.measen.2022.100482
24. Dharmarajan, K., et al.: Thyroid disease classification using decision tree and SVM. Indian J. Public Health Res. & Develop. 11(3), 224–229 (2020)
25. Abbad Ur Rehman, H., et al.: Performance analysis of machine learning algorithms for thyroid disease. Arabian J. Sci. Eng. 46(10), 1–13 (2021). https://doi.org/10.1007/s13369-020-05206-x
26. Raghuraman, M., Sailatha, E., Gunasekaran, S.: Efficient thyroid disease prediction and comparative study using machine learning algorithms. Int. J. Inf. Comput. Sci. 6(6), 617–624 (2019)
27. Sultana, A., Islam, R.: Machine learning framework with feature selection approaches for thyroid disease classification and associated risk factors identification. J. Electr. Syst. Inf. Technol. 10(1), 1–23 (2023). https://doi.org/10.1186/s43067-023-00101-5
28. Yadav, D.C., Pal, S.: Decision tree ensemble techniques to predict thyroid disease. Int. J. Recent Technol. Eng. 8(3), 8242–8246 (2019). https://doi.org/10.35940/ijrte.c6727.098319
29. Yadav, D.C., Pal, S.: Prediction of thyroid disease using decision tree ensemble method. Human-Intelligent Syst. Integration 2(1-4), 89–95 (2020). https://doi.org/10.1007/s42454-020-00006-y
30. Priya, S.D., et al.: Diagnostics of Thyroid using Support Vector Machine
31. Shen, Y., et al.: Diagnosis of thyroid neoplasm using support vector machine algorithms based on platelet RNA-seq. Endocrine 72(3), 758–783 (2021). https://doi.org/10.1007/s12020-020-02523-x
32. Mir, Y.I., Mittal, S.: Thyroid disease prediction using hybrid machine learning techniques: an effective framework. Int. J. Scientific & Technol. Res. 9(2), 2868–2874 (2020)
33. Hosseinzadeh, M., et al.: A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things. J. Supercomput. 77(4), 3616–3637 (2021). https://doi.org/10.1007/s11227-020-03404-w
34. Sharifi, A., Alizadeh, K.: Comparison of the particle swarm optimization with the genetic algorithms as a training for multilayer perceptron technique to diagnose thyroid functional disease. Shiraz E-Medical J. 22(1) (2021). https://doi.org/10.5812/semj.100351
35. Jha, R., Bhattacharjee, V., Mustafi, A.: Increasing the prediction accuracy for thyroid disease: a step towards better health for society. Wireless Pers. Commun. 122(2), 1921–1938 (2022). https://doi.org/10.1007/s11277-021-08974-3
36. Sankar, S., et al.: Thyroid disease prediction using XGBoost algorithms. J. Mob. Multimed 18, 1–18 (2022). https://doi.org/10.13052/jmm1550-4646.18322
37. Sureshkumar, V., et al.: A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine. Expet Syst. 39(1), e12811 (2022). https://doi.org/10.1111/exsy.12811
38. Kumar, S., et al.: Butterfly optimized feature selection with fuzzy C-means classifier for thyroid prediction. Intell. Automation & Soft Comput. 35(3), 2909–2924 (2023). https://doi.org/10.32604/iasc.2023.030335
39. Shankar, K., et al.: Optimal feature-based multi-kernel SVM approach for thyroid disease classification. J. Supercomput. 76(2), 1128–1143 (2020). https://doi.org/10.1007/s11227-018-2469-4
40. Pradeep, K., et al.: Improved machine learning method for intracranial tumor detection with accelerated particle swarm optimization. J. Healthc. Eng. 2022, 1–13 (2022). https://doi.org/10.1155/2022/1128217
41. Umar Sidiq, D., Aaqib, S.M., Khan, R.A.J.I.J.S.R.C.S.I.T.: Diagnosis of various thyroid ailments using data mining classification techniques. Int. J. Scientific Res. Comput. Sci. Eng. Inf. Technol. 5, 131–136 (2019). https://doi.org/10.32628/cseit195119
42. Zhou, M.-G., et al.: Quantum neural network for quantum neural computing. Research 6, 0134 (2023). https://doi.org/10.34133/research.0134

How to cite this article: Sha, M.: Quantum intelligence in medicine: empowering thyroid disease prediction through advanced machine learning. IET Quant. Comm. 5(2), 123–139 (2024). https://doi.org/10.1049/qtc2.12078