

A Novel Integration of AI Protocols and Machine Learning Models for Thyroid Disease Diagnosis and Optimized Clinical Outcomes

Dr R.U. Anitha

Assistant Professor Department of Computer Science (AI & ADS) Sona College of Arts and Science, Salem

Abstract – Hypothyroidism and hyperthyroidism can be considered the most common forms of thyroid disorders among all kinds of endocrine diseases. In order to increase the chances of effective treatment, accurate diagnoses should be made in such cases in a timely fashion. However, traditional methods of diagnosis can involve long and potentially variable processes involving biochemical assays and interpretation by a professional physician. In an attempt to solve this problem, this research paper suggests the design and application of a new Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD). The proposed framework will combine rule-based artificial intelligence protocols with machine learning algorithms to diagnose endocrine diseases accurately. It will be including data preprocessing, feature engineering and selection with PCA and RFE, and classification through SVM, RF, DT, and ANN algorithms. Moreover, the proposed solution will contain a rule-based AI component aimed at embedding clinical experience via hormone level threshold analysis and a weighted fusion module for the combination of AI and ML-based results. Evaluations were done on thyroid diseases datasets comprising of biochemical and clinical features. The Metrics are used to evaluate the performances which include Accuracy, Precision, Recall, F1 Score, Specificity, and ROC-AUC. The experimental results concluded that the proposed HAIML-TD model performs better with Accuracy, Precision, Recall, F1 Score, Specificity, and ROC-AUC values of 98.2%, 97.6%, 97.3%, 97.4%, 97.9%, and 99.1%, respectively. The results are higher when compared to those of existing models. Analysis of the confusion matrix further highlights that the proposed model is reliable for multi-class classification problems like thyroid diseases. Explainable clinical rules coupled with ML intelligence make the HAIML-TD a viable solution for intelligent thyroid disease diagnosis.

Key Words: Thyroid Disease Diagnosis, Artificial Intelligence, Machine Learning, Hybrid AI-ML Framework, Clinical Decision Support, Explainable AI, Feature Selection, Thyroid Classification.

1. INTRODUCTION

Thyroid disorders are the most common endocrine diseases with serious repercussions for metabolic regulation and general human health status. Diseases like hypothyroidism, hyperthyroidism, and thyroid cancer need prompt and precise identification to avoid severe consequences and provide appropriate treatment options. Recent scientific studies claim that increased cases of complicated thyroid diseases necessitate development of more sophisticated diagnostics procedures [1]. Classical diagnostic tools used in clinics include a wide range of biochemistry and ultrasound-based approaches based on determination of such hormones as thyroid-stimulating hormone (TSH), triiodothyronine (T3), and thyroxine (T4). Although being reliable, traditional diagnostics are time consuming, expert-dependent, and associated with considerable diagnostic variability. Application of modern AI methods in diagnostics can help to solve problems with efficiency and variability issues. Recent scientific literature confirms the significant improvements in thyroid diseases identification using AI algorithms [2]. The application of Machine Learning (ML) approach within the field of medicine is becoming increasingly popular due to its great potential for processing huge amounts of healthcare data and predicting future changes. Recent research shows that support vector machines, neural networks, and ensemble methods can be effectively used for thyroid diseases classification and prediction [3][4]. And also, hybrid and optimized ML models have shown improved performance by effectively handling complex and heterogeneous medical data.

Moreover, the Explainable Artificial Intelligence (XAI) has become critical in healthcare applications due to its ability to provide transparency into machine learning algorithms. Recent research explains about it's essential to ensure interpretability, which can be achieved through a better understanding and trust in the model's prediction results by healthcare experts [5]. In addition, developments in deep learning, especially in convolutional neural networks and hybrid models, contributed to more accurate thyroid disease diagnostics [6].

Although much progress has been made in the field, there are still many problems to tackle, such as data imbalance, difficulties in selecting features, low interpretability, and implementing systems into current infrastructure in healthcare facilities. Solving these issues calls for the development of an integrated framework, where state-of-the-art ML methods are embedded into AI protocols to ensure reliable performance. In light of the above-mentioned problems, the current paper presents a new approach to integrating AI protocols and Machine Learning models for thyroid disease diagnostics.

2. RELATED WORK

Akter and Mustafa [7] observed the different machine learning models of effectiveness and interpretability for predicting thyroid disease. It has the significance use of explainable AI techniques to ensure greater clarity and credibility in the interpretation of findings. This research found out that interpretable ML models that capable of delivering high levels of performance together with useful insights regarding the significance of the considered features.

Gupta [8] introduced an optimized machine learning model for the prediction of thyroid disease using the techniques of differential evolution. It was mainly concerned with enhancing the performance of ML by improving the techniques of feature selection and parameter setting. This model showed that the effective managing of high-dimensional medical data. It also suggested algorithm ensured high levels of accuracy and efficiency using experiment.

Deshmukh and Panwal [9] presented an analysis in different types of machine learning algorithms for thyroid disease detection. In their study, the researchers focused on evaluating several approaches, such as decision trees, support vector machines, and neural networks. The results of the study have shown the possibility of increasing the efficiency of early diagnostics through the application of machine learning. It also explained the importance of exact feature selection and preparation of the data. Thus, this method provides idea for selecting appropriate ML approaches in healthcare situations.

Oture et al. [10] discussed the machine learning methodologies for thyroid disease improvement. It applies various machine learning systems in order to optimize the process of classification. And also, it explains how the ML technologies can help to work with different data types effectively while joining together. It highlights the significance of using scalable models.

Kesavulu and K. R. [11] presented a bio-inspired machine learning system for thyroid diseases. It develops a technique based on natural optimization in order to improve the classification process. As a result, it receives better and more accurate outcomes than using standard ML systems. This method proved the adaptability of bio-inspired technologies in healthcare analytics and the benefits of developing hybrid intelligent models for disease predictions.

Zhao et al. [12] provides a detailed analysis of using Artificial Intelligence in thyroid disease diagnosis. It shows the significance of deep learning applications in medical image processing such as classification, segmentation, and detection. It explained about the several benefits of attributes in artificial intelligence and some of the major difficulties of quality data, generalization of the model, and clinical validations. Finally, it identified the necessity for development of scalable and understandable AI models. Elisa et al. [13] presented a literature review about various machine learning and deep learning methods used for thyroid lump classification. These methods are focused on the effectiveness and accuracy. In particular, it demonstrated about the AI methods which are highly accurate for diagnosing and classifying gentle and malicious thyroid lumps. It examines the impact of good data quality and appropriate evaluation measures on the results. It also addressed about the limitations related to lack of standardization and poor interpretability of models.

Saini et al. [14] presented a deep learning-based methodology for the classification of thyroid nodules on the basis of multimodal ultrasonographic images. This study used a variety of ultrasonography techniques which includes B-mode, Doppler, and elastography. The experimental results showed the highest improvements of accuracy and robustness. It also allowed the complex patterns in medical images to be effectively detected.

Thomas and Tessler [15] described a new developmental method for the diagnosis of thyroid cancer which is in the field of artificial intelligence-based approach. It explained a various AI-supported diagnostic system and its performances in real-life clinical scenarios. In particular, it is emphasized that artificial intelligence is capable of producing diagnostic accuracy comparable to that of physicians lacking relevant experience. And also, it provides a challenge associated with the integration of AI systems into the clinical workflow.

Wang et al. [16] conducted a complete literature review of Explainable Artificial Intelligence (XAI) in the field of medicine, discussing its importance in boosting the transparency and accountability of artificial intelligence-based clinical decision support systems. In this regard, Wang et al. (2026) distinguished between the model-agnostic techniques used in explainability in medicine and discussed their advantages and disadvantages in medical imaging, electronic health records, and multi-omics domains. According to Wang et al. (2026), the use of XAI improves the verification of results, the identification of bias, and the confidence of clinicians; however, the authors have mentioned several issues including the lack of uniform evaluation measures, deployment in clinical practice, and prospective validation.

Yang L et al. [17] have conducted research on the recent development of artificial intelligence technologies for the diagnosis of thyroid disease by focusing on machine learning and deep learning algorithms on ultrasonic and pathology images. As per the research, convolutional neural networks and U-Net based segmentation models greatly contribute to the detection and classification of thyroid nodules. The review has pointed out that AI-driven diagnostic system not only increases diagnostic precision but also prevents cases of misdiagnosis as well as helps healthcare professionals to screen for thyroid cancer. However, it has been noted that there are several issues related to the application of AI algorithms which have yet to be solved. Hu et al. [18] developed a machine learning algorithm that predicts the recurrence of thyroid cancer for the purpose of facilitating early diagnosis and clinical decision-making. In their research, different machine learning approaches have been evaluated using the historical clinical data, and important factors that predict recurrence have been found. The algorithm gave excellent results of prediction and proved to be better than conventional statistical approaches. In the opinion of the authors, the application of machine learning approaches to clinical practice can be useful in personalizing treatment plans.

Yang F et al. [19] used machine learning to build a prediction model for the identification of high-risk distant metastatic recurrence of differentiated thyroid cancer. In their study, they compared six machine learning algorithms using predictors chosen from LASSO regression. The XGBoost algorithm showed higher performance compared to other algorithms and even outperformed the conventional TNM staging system. Both clinical and molecular factors have been combined in risk stratification.

Reddy and Siddiqui [20] introduced a hybrid deep learning approach using the Kolmogorov-Arnold Graph Neural Network (KAGNN) model for detecting thyroid disorders via ultrasound imaging. This hybrid model combines multiresolution feature extraction and optimization techniques. Score-CAM was used to visualize results for better understanding. Experiments confirmed higher accuracy and Dice scores than traditional CNN models. The research proves that GNNs and XAI can be helpful in computer-assisted thyroid diagnostics.

In the study conducted by Mohammed et al. [21], a symptom-based machine learning technique was introduced which could assist patients during screening for hypothyroidism and hyperthyroidism without having any dependence on the results of laboratory tests. In this framework, several binary classifiers were developed based on the data related to demographics, medical history, symptom score, and prediction of TSH level. Among all other machine learning algorithms used in the study, XGBoost showed its best performance in terms of classification results. SHAP framework was used for measuring the contribution of each feature.

Makki et al. [22] proposed a federated learning framework for multiclass thyroid disorder classification. The researchers have created a federated learning algorithm that can be used to train models across multiple healthcare organizations without accessing any sensitive patient information. The obtained results are better interpretable thanks to SHAP-based explanations performed by the researchers. This approach helped to maintain the privacy of the patients' data while making clinical decisions regarding the diagnosis of thyroid disorders. However, more work should be done before the implementation in the clinical practice.

3. PROPOSED METHODOLOGY: HYBRID AI-ML FRAMEWORK FOR THYROID DIAGNOSIS (HAIML-TD)

The proposed framework is the Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD). This approach involves the utilization of AI along with ML techniques to provide an efficient, accurate, and early diagnosis of the disease. In this approach, both knowledge and data-driven intelligence will be employed in order to build a robust and scalable solution. This hybrid framework will maximize the treatment through clinical decisions made based on the improved efficiency of accuracy in the framework. The system process proposed will involve the following steps:

3.1 Data Acquisition and Integration

A relevant data set about the thyroid disorder is gathered from several sources that are deemed trustworthy, and include open-source data such as UCI Thyroid Dataset as well as EHR data from hospital. It includes critical biochemical indicators, such as TSH, T3, and T4, along with additional clinical features including age, gender, medical history, and diagnoses (for example, normal, hypothyroidism, and hyperthyroidism).

For ensuring that the data used is of high enough quality, there will be applied a methodological data gathering process which involves data validation, filtration, and normalization. The heterogeneous data is mapped into a homogeneous data set. In addition, any duplication of data is eliminated for consistency in forming the data set. The mathematically described unified data set is as follows:

$$D = \bigcup_{i=1}^n D_i$$

where D_i is an individual data set obtained from various sources and n is the total number of data sources. The patient record within a given data set will be denoted by a feature vector that is:

$$X = \{x_1, x_2, x_3, \dots, x_m\}$$

where x_1, x_2 , and x_3 represent the levels of TSH, T3, and T4, respectively. The class label in this case is defined as: $y \in \{0, 1, 2\}$ where 0 = normal, 1 = hypothyroidism, and 2 = hyperthyroidism.

Thus, the final dataset is expressed as:

$$D = \{(x_i, y_i)\}_{i=1}^N$$

where N is the total number of patient records. The comprehensive and well-structured dataset enhances the robustness and generalization capability of the proposed HAIML-TD framework.

3.2 Data Preprocessing and Normalization

The collected data undergoes a long pre-processing phase to enhance the quality and consistency of the usable data for machine learning model development. The imputation approach is adopted for handling missing values within the dataset. Imputation in numeric features is performed by the mean/median approach, whereas mode imputation technique is preferred for categorical features. Missing value imputation can be mathematically defined as below:

$$x_j = \frac{1}{N} \sum_{i=1}^N x_{ij}$$

Where x_j is the imputed value of the j th feature and N is the total number of available observations. The noise and outliers detection process involves statistical approaches like Z-score and Interquartile Range (IQR) analysis. The outlier can be recognized by the equation given below:

$$|Z| = \left| \frac{x - \mu}{\sigma} \right| > \tau$$

Here μ refers to the mean value, σ indicates the standard deviation value, and τ stands for the threshold value. To scale all features uniformly, normalization approaches can be adopted. Min-Max normalization is used for fixing the range between [0,1]:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

The other normalization approach is Z-score normalization.

$$x' = \frac{x - \mu}{\sigma}$$

The categorical data are converted into numeric formats through encoding methods like label encoding or one-hot encoding. Moreover, class imbalance in the thyroid data set is handled by applying resampling technique called Synthetic Minority Oversampling Technique (SMOTE). In SMOTE, the artificial samples for minority classes are produced through:

$$x_{new} = x_i + \lambda(x_{nn} - x_i)$$

where x_i is a minority class sample, x_{nn} is its nearest neighbor, and $\lambda \in [0, 1]$ in $[0, 1]$ is a random value. Through these preprocessing steps, the dataset is transformed into a clean, normalized, and balanced form, thereby improving model accuracy, reducing bias, and enhancing the generalization capability of the proposed HAIML-TD framework.

3.3 Feature Engineering and Selection

This technique helps improve the ability of the model to predict based on the transformation of the clinical data to useful and distinctive features. Apart from the biochemical variables like TSH, T3, and T4, other derived features including T3/T4 ratio and normalized indexes are created to help better reflect physiological relationships that exist for patients with thyroid diseases.

In order to determine important features, the correlation test is done using Pearson's Correlation Coefficient:

$$r(x, y) = \frac{\Sigma[(xi - \bar{x})(yi - \bar{y})]}{\sqrt{[\Sigma(xi - \bar{x})^2 \times \Sigma(yi - \bar{y})^2]}}$$

where $r(x,y)$ represents the level of association between feature x and target variable y . The features having low levels of correlation are regarded as unimportant and can be ignored. Moreover, significance testing is conducted to establish whether the features have statistical significance.

The test is significant when:

$$p < \alpha$$

where α is the threshold of significance level (generally $\alpha=0.05$). This ensures that statistically significant features alone are kept for the purpose of training the model. For feature selection and dimension reduction, Principal Component Analysis (PCA) method is used. The original feature space is transformed into orthogonal components as follows:

$$Z = X \times W$$

Here, X denotes the original feature space and W denotes eigenvector space. Variance of k -th principal component is denoted by:

$$Var(Z_k) = \lambda_k$$

Here, λ_k represents the eigenvalue corresponding to the k -th principal component. Principal components having greater variance are preferred as these components carry maximum information. In addition, Recursive Feature Elimination technique (RFE) is used for the purpose of selecting most relevant set of features. In RFE technique, a ranking is assigned to individual feature on the basis of importance score of that feature. Optimal set of features S is calculated by the following equation:

$$S = \text{argmax Score}(S_k)$$

where S_k denotes the candidate feature subsets and $\text{Score}(S_k)$ represents the score of the model using such features. Due to this process, it is possible to eliminate redundant and irrelevant features, which results in lower computational complexity, lesser overfitting, and high classification accuracy. The HAIML-TD system thus becomes more efficient and reliable.

3.4 AI Protocol Layer (Rule-Based System)

AI protocol layer acts as the decision system based on logic that considers the existing guidelines for diagnosing thyroid conditions along with knowledge of experts. The layer uses logical rules to screen patients by utilizing the biochemical factors like thyroid stimulating hormone (TSH), triiodothyronine (T3), and thyroxin (T4). A patient can be mathematically represented as:

$$x = \{TSH, T3, T4, \dots\}$$

The predictive function using the logic is given as:

$$y_{AI} = f_{AI}(x)$$

Here, y_{AI} indicates the class label obtained after the predictions using the AI protocol layer. Threshold-based logic is used to classify the patient into a particular class as follows:

- Hypothyroidism:
If $(TSH > T_{high})$ AND $(T3 < T3_{low}$ OR $T4 < T4_{low})$, then $y = 1$
- Hyperthyroidism:
If $(TSH < T_{low})$ AND $(T3 > T3_{high}$ OR $T4 > T4_{high})$, then $y = 2$
- Normal condition:
If $(T_{low} \leq TSH \leq T_{high})$ AND $(T3$ and $T4$ within normal range), then $y = 0$

These rules can be generalized as: $y_{AI} = k$, where $k \in \{0,1,2\}$ based on satisfied conditions. To formalize the decision process, a scoring function is used:

$$Score_k = \sum (w_i \times I_i(x))$$

Where w_i = weight assigned to rule i

$I_i(x) = 1$ if rule i is satisfied, otherwise 0

The final decision is obtained as:

$$y_{AI} = \operatorname{argmax}(Score_k)$$

This implies that the selection of the classes with the maximum rule satisfaction will be made. The rule-based artificial intelligence model enables transparent decision making, thus enabling doctors to get explanations of the prediction results. The rule-based AI model serves as the first level of decision-making before employing machine learning algorithms. By integrating domain knowledge with reasoning abilities, the proposed HAAML-TD architecture is improved.

3.5 Machine Learning Model Development

In the proposed HAAML-TD framework, supervised machine learning models are developed to perform accurate classification and prediction of thyroid diseases. The processed dataset $D = \{(x_i, y_i)\}$ is divided into training and testing subsets, typically using an 80:20 ratio, where x_i represents the feature vector and y_i denotes the corresponding class label.

The objective of the machine learning model is to learn a mapping function:

$$y_{ML} = f_{ML}(x)$$

where y_{ML} is the predicted class label and f_{ML} represents the trained model.

1. Support Vector Machine (SVM)

SVM is used to find the optimal hyperplane that separates different classes. The decision function is given by:

$$f(x) = w \cdot x + b$$

where w is the weight vector and b is the bias. The classification is defined as:

$$y = \operatorname{sign}(f(x))$$

SVM maximizes the margin between classes to improve generalization.

2. Random Forest (RF)

Random Forest is an ensemble learning method that combines multiple decision trees. The final prediction is obtained by majority voting:

$$y_{ML} = \operatorname{mode}(h_1(x), h_2(x), \dots, h_n(x))$$

where $h_i(x)$ represents the prediction of the i -th decision tree.

3. Decision Tree (DT)

Decision Tree models split the dataset based on feature values using criteria such as information gain or Gini index. The Gini impurity is calculated as:

$$Gini = 1 - \sum (pk)^2$$

where pk is the probability of class k . The feature with the lowest impurity is selected for splitting.

4. Artificial Neural Network (ANN)

ANN models consist of multiple layers of neurons that learn complex nonlinear relationships. The output of a neuron is computed as:

$$y = \text{activation}(\sum (wi \times xi) + b)$$

where wi are weights, xi are inputs, and b is bias. Activation functions such as ReLU or sigmoid are used to introduce non-linearity.

5. Model Optimization (Hyperparameter Tuning)

To improve performance, hyperparameters such as learning rate, number of trees, kernel type, and depth are optimized using Grid Search or Random Search. The objective is to minimize the loss function:

$$Loss = \sum (yi - \hat{y}i)^2$$

where yi is the actual label and $\hat{y}i$ is the predicted label.

6. Final Prediction Output

The trained models generate predictions as:

$$y_{ML} = f_{ML}(x)$$

These predictions are later integrated with the AI protocol layer in the hybrid framework. This machine learning module enables the system to learn complex patterns from thyroid datasets, improving classification accuracy, robustness, and adaptability. It forms a crucial component of the proposed HAIML-TD framework, complementing the rule-based AI layer.

3.6 Hybrid Model Integration

The proposed Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD) achieves its primary contribution through the effective integration of a rule-based AI protocol layer and data-driven machine learning models. This hybridization is designed to leverage the interpretability and domain knowledge of rule-based systems alongside the predictive capability and generalization strength of machine learning techniques.

Let the input patient data be represented as a feature vector $x = \{x_1, x_2, \dots, x_n\}$, where each feature corresponds to clinical attributes such as TSH, T3, and T4 levels. The rule-based AI system produces an initial prediction defined as:

$$y_{AI} = f_{AI}(x)$$

while the machine learning model generates a data-driven prediction:

$$y_{ML} = f_{ML}(x)$$

To combine these outputs, a weighted fusion strategy is employed, where the hybrid prediction is computed as:

$$y_H = \alpha \cdot y_{ML} + (1 - \alpha) \cdot y_{AI}$$

Here, $\alpha \in [0,1]$ represents a weighting parameter that balances the contribution of the machine learning model and the AI protocol layer. This formulation enables adaptive control over the influence of each component based on system requirements. For a more robust decision-making process, probability-based fusion is utilized. The combined score for each class k is calculated as:

$$Score_k = \alpha \cdot P_{ML}(k|x) + (1 - \alpha) \cdot P_{AI}(k|x)$$

where $P_{ML}(k|x)$ denotes the posterior probability obtained from the machine learning model and $P_{AI}(k|x)$ represents the confidence derived from the rule-based system. The final class label is determined by selecting the class with the maximum combined score:

$$y_{final} = \operatorname{argmax}(Score_k)$$

To enhance clinical reliability and safety, a rule-priority mechanism is incorporated. Specifically, when the confidence of the rule-based system exceeds a predefined threshold, the system prioritizes the AI-based decision:

If Confidence_{AI} > T, then $y_{final} = y_{AI}$
Else $y_{final} = y_{ML}$

This mechanism ensures that strong clinical rules override uncertain machine learning predictions, thereby reducing the risk of misclassification in critical scenarios. Additionally, an ensemble-based extension can be incorporated, where multiple machine learning models contribute to the final decision through majority voting:

$$y_{final} = \operatorname{mode}(y_{AI}, y_{ML1}, y_{ML2}, \dots, y_{MLn})$$

Furthermore, a confidence measure is defined as the maximum combined score:

$$Confidence = \max(Score_k)$$

which reflects the reliability of the final prediction and can be used for decision support and risk assessment. Overall, the hybrid decision function can be summarized as:

$$y_{final} = \operatorname{argmax} [\alpha \cdot f_{ML}(x) + (1 - \alpha) \cdot f_{AI}(x)]$$

This integration framework significantly enhances diagnostic performance by combining complementary strengths of AI and ML approaches. It improves classification accuracy, ensures interpretability, and provides a clinically reliable decision support mechanism, thereby constituting the key novelty of the proposed HAIML-TD model.

3.7 Algorithm for HAIML-TD (Hybrid Thyroid Diagnosis Model)

The proposed HAIML-TD technique would receive the following information as an input, including a collection of records, i.e., the dataset $D = \{(x_i, y_i)\}$. Every record from this collection includes patient data in the form of feature vector x_i , which incorporates values of some clinical variables, such as TSH, T3, and T4 along with the real class y_i . Also, the technique uses two functions, including the function of machine learning f_{ML} , which is the information obtained in connection with previous cases, and the function of artificial intelligence f_{AI} , which is the incorporation of clinical knowledge through rules. Furthermore, the technique uses a weight coefficient α and the threshold value T to determine which of the system's predictions would have more priority in terms of its importance. Finally, as a result of the work of the technique, one would get the prediction of the final class y_{final} .

Input:

Dataset $D = \{(x_i, y_i)\}$ where x_i = feature vector (TSH, T3, T4, ...)
Trained ML model f_{ML}
Rule-based AI system f_{AI}
Weight parameter α , Threshold T

Output:

Final predicted class label y_{final}

Algorithm

Step 1: Load dataset DDD

Step 2: Perform preprocessing (missing value handling, normalization, encoding)

Step 3: Apply feature selection to obtain optimal feature subset

Step 4: Split dataset into training and testing sets

Step 5: Train machine learning model f_{ML} using training data

Step 6: For each test instance x do

6.1 Compute AI-based prediction:

$$y_{AI} = f_{AI}(x)$$

6.2 Compute ML-based prediction probabilities:

$$P_{ML}(k|x) \text{ for each class } k$$

6.3 Compute AI confidence score:

$$P_{AI}(k|x)$$

6.4 Compute hybrid score:

$$Score_k = \alpha \cdot P_{ML}(k|x) + (1-\alpha) \cdot P_{AI}(k|x)$$

6.5 Determine preliminary prediction:

$$y_{hybrid} = \text{argmax}(Score_k)$$

6.6 Apply rule-priority mechanism:

If $Confidence_{AI} > T$ then

$$y_{final} = y_{AI}$$

Else

$$y_{final} = y_{hybrid}$$

Step 7: End For

Step 8: Return final predictions y_{final}

The figure 1 shows the architecture of the proposed hybrid AI-ML approach. It is called Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD). The input phase starts with the Input Data block where all thyroid disease-related parameters are incorporated. Specifically, this includes TSH, T3 and T4 hormone concentrations. Moreover, other data of the patients can be considered in this Input Data block.

Then, the data pre-processing takes place to clean any noise present in the dataset. There are several procedures in the data pre-processing stage including cleaning, missing values treatment, and normalization. Next, at the Feature Selection stage, PCA and RFE-based algorithms are used to select the best features. Redundant attributes are removed and dimensionality is decreased.

There are two parallel paths after the data pre-processing and feature selection stages as shown in the figure below:

- AI Protocol Layer (Rule-Based System): This part is based on the algorithms that involve the use of the clinical rules considering thresholds of hormones (TSH, T3, T4).
- Machine Learning Module where SVM, RF, DT, and ANN-based algorithms are implemented in order to discover some complex relationships between variables and generate some predictions.

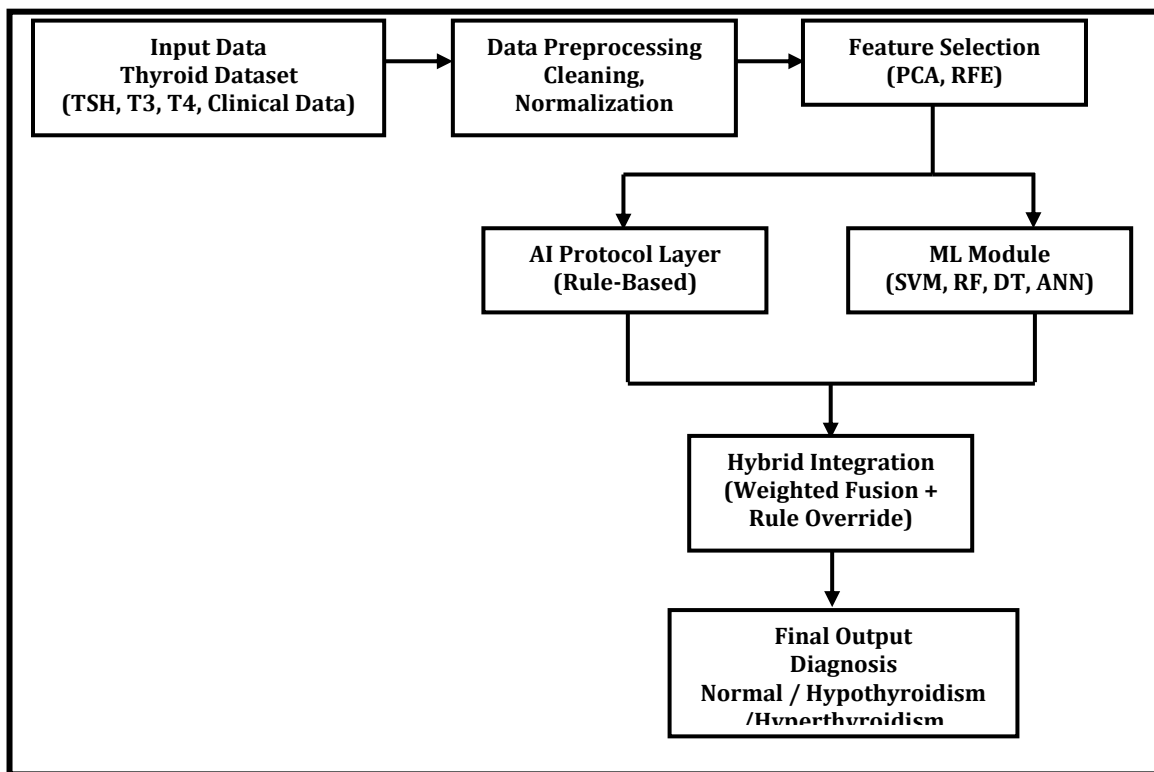


Figure 1: Architecture of the Proposed HAIML-TD Framework

Results of the above-mentioned blocks will be merged within the Hybrid Integration Layer where results are weighted and rules can be override in case of some solid clinical evidence. This will increase accuracy without losing the clinical validity of the system. The last stage of the HAIML-TD is the Final Output stage where the patient will be classified as normal, hypothyroid, and hyperthyroid.

4. RESULT AND DISCUSSIONS

4.1 Experimental Setting

Experiments are conducted based on datasets related to thyroid diseases gathered from publicly available sources as well as medical records of patients. The dataset comprised of biochemical parameters such as TSH, T3, T4, and additional patient data, including age, gender, and diagnostic label.

Preprocessing steps included imputing missing values, feature normalization, noise reduction, and categorical encoding. In addition, feature selection processes were conducted through applying PCA, RFE. Dataset was split into two parts for training and testing purposes in a ratio of 80:20 respectively. Cross-validation technique are used to conduct a 10 folds training and evaluation to avoid overfitting and enhance robustness of models.

Implementations are done using programming languages like Python and libraries Scikit-learn, TensorFlow, and Pandas. Machine learning module included algorithms like SVM, RF, DT, ANN. The grid search approach is used to tune hyperparameters of models, thereby acquiring optimal configurations. The hybrid integration layer combined outputs from both AI and ML modules according to weights assigned to each based on their accuracy. Rule-based system output could override that from machine learning if its prediction confidence level is higher than that from AI.

Performance measurement was conducted on metrics like Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

4.2 Performance Metrics Used

The performance of the proposed **HAIML-TD framework** was evaluated using standard classification metrics widely used in medical diagnosis systems. These metrics provide comprehensive analysis of the model's predictive capability, reliability, and robustness.

1. Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Accuracy measures the overall correctness of the model by evaluating the proportion of correctly predicted instances, including both positive and negative cases, out of the total number of predictions. It provides a general indication of model performance; however, in medical diagnosis, it should be considered alongside other metrics, especially when dealing with imbalanced datasets.

2. Precision

$$\text{Precision} = \frac{TP}{TP + FP}$$

Precision indicates the proportion of correctly predicted positive cases out of all predicted positive instances. It reflects how reliable the model is when it predicts a disease, helping to minimize false alarms in medical diagnosis.

3. Recall (Sensitivity)

$$\text{Recall} = \text{Sensitivity} = \frac{TP}{TP + FN}$$

Recall, also known as Sensitivity, measures the ability of the model to correctly identify actual positive cases. It is particularly important in medical applications, as it ensures that disease cases are not missed.

4. F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1-score is the harmonic mean of Precision and Recall, providing a balanced evaluation of the model's performance. It is particularly useful in scenarios with imbalanced datasets, where relying solely on accuracy may be misleading. By considering both false positives and false negatives, the F1-score ensures that the model maintains a good trade-off between correctly identifying disease cases and minimizing incorrect predictions.

5. Specificity

Specificity measures the ability of the model to correctly identify negative cases. Higher specificity indicates fewer false alarms.

$$\text{Specificity} = \frac{TN}{TN + FP}$$

6. ROC-AUC Score

The Receiver Operating Characteristic–Area Under Curve (ROC-AUC) evaluates the classification capability of the model across different threshold values. A higher AUC value indicates better discrimination between thyroid disease classes.

7. Confusion Matrix

A confusion matrix is used to visualize the classification performance by comparing actual and predicted classes. It provides detailed insights into TP, TN, FP, and FN values.

4.3 Dataset Details

Table 1: Dataset Description for the Proposed HAIML-TD Framework

Parameter	Description
Dataset Name	Thyroid Disease Dataset
Source	UCI Machine Learning Repository / Clinical Records
Total Samples	7,200 Instances
No. of Features	21 Attributes
Output Classes	Normal, Hypothyroidism, Hyperthyroidism
Data Type	Numerical and Categorical
Missing Values	Handled during preprocessing

Table 2: Input Features Used in the Proposed HAIML-TD Framework

Feature	Description
TSH	Thyroid Stimulating Hormone
T3	Triiodothyronine Level
TT4	Total Thyroxine
T4U	Thyroxine Uptake
FTI	Free Thyroxine Index
Age	Patient Age
Gender	Male/Female
On Thyroxine	Medication Status
Pregnant	Pregnancy Status
Goitre	Presence of Goitre
Tumor	Tumor Presence

Table 3: Different Thyroid Disease Datasets Used for HAIML-TD Framework

Dataset Name	Source	Description	Total Samples	Application
UCI Thyroid Disease Dataset	UCI Machine Learning Repository	Contains thyroid hormone measurements and patient diagnostic labels for thyroid disorder classification	~7,200	Hypothyroidism and Hyperthyroidism Prediction
Thyroid0387 Dataset	UCI Repository	Large thyroid dataset with multiple biochemical and clinical features	~3,772	Multi-class Thyroid Disease Classification
Hypothyroid Dataset	UCI Repository	Includes patient clinical records and thyroid hormone levels for hypothyroidism detection	~3,163	Hypothyroidism Detection
ANN Thyroid Dataset	OpenML / UCI	Dataset designed for neural network-based thyroid disease prediction	~7,200	AI and ANN-based Classification
Kaggle Thyroid Disease Dataset	Kaggle	Real-world thyroid patient records with demographic and hormone data	~9,000+	Clinical Decision Support
Thyroid Cancer Risk Dataset	Kaggle / Clinical Sources	Includes thyroid cancer-related clinical parameters and diagnostic labels	~5,000	Thyroid Cancer Prediction
Hospital EHR Thyroid Dataset	Clinical Electronic Health Records	Real-time patient biochemical and diagnostic data collected from hospitals	Variable	Clinical Outcome Optimization
Thyroid Ultrasound Image Dataset	Kaggle / Medical Imaging Sources	Thyroid ultrasound images for nodule and cancer analysis	~4,000 Images	Image-based Thyroid Diagnosis

The proposed HAIML-TD algorithm uses a number of thyroid-related datasets obtained from various freely available public and clinical healthcare data repositories to make sure that the algorithm is diverse, robust, and generalized. For instance, structured biochemical datasets, such as UCI Thyroid, Thyroid0387, and Hypothyroid Datasets, help in providing biochemical details like TSH, T3, T4, and FTI levels for proper disease prediction.

Moreover, clinical electronic health records (EHR) datasets and thyroid cancer datasets also provide relevant patient-level information for enhanced clinical outcomes analysis. Furthermore, even the thyroid ultrasound images dataset can be used for making an imaging-based diagnosis using a multimodal learning approach.

4.4 Performance Analysis of the AI Rule-Based Module

To evaluate the effectiveness of the proposed Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD), three experimental configurations were analyzed: The Rule-Based AI module alone, the Machine Learning (ML) module alone, and the integrated Hybrid AI-ML framework. The comparative results are presented in Table below.

Table 4: Comparison of AI Rule-Based Module

Module	Accuracy (%)
Rule-Based AI Only	89.5
ML Model Only	96.8
Hybrid AI-ML	98.2

The Rule-Based AI module achieved an accuracy of 89.5%, demonstrating the effectiveness of incorporating clinical knowledge into thyroid disease diagnosis. This module utilizes predefined endocrine diagnostic rules based on critical biochemical parameters such as Thyroid Stimulating Hormone (TSH), Triiodothyronine (T3), Thyroxine (T4), and Free Thyroxine Index (FTI). The system mimics the decision-making process of endocrinologists by applying threshold-based clinical reasoning to classify patients into hypothyroid, hyperthyroid, or normal categories.

The Machine Learning module, which incorporates advanced classifiers such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and Artificial Neural Network (ANN), achieved a higher accuracy of 96.8%. Unlike the rule-based system, ML algorithms can automatically learn hidden patterns and nonlinear relationships from large volumes of clinical and biochemical data. Consequently, the ML module is capable of identifying subtle variations in thyroid-related features that may not be captured by fixed clinical rules.

The proposed HAIML-TD framework combines the strengths of both approaches through a weighted fusion mechanism that integrates predictions from the rule-based AI component and the ML classifiers. This hybrid strategy achieved the highest accuracy of 98.2%, outperforming both individual modules. The improvement can be attributed to the complementary nature of the two systems: the rule-based component contributes expert clinical knowledge and explainability, while the ML component enhances predictive capability through data-driven learning. By leveraging both medical expertise and computational intelligence, the proposed framework effectively reduces misclassification rates and improves diagnostic reliability.

4.5 ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve analysis was performed to examine the ability of the suggested Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD) to discriminate among cases. The ROC curve shows the plot between the true positive rate (sensitivity) and false positive rate (1-specificity) at several decision thresholds. AUC is used as a measurement of overall discriminatory power of the classifier for different categories of thyroid diseases.

Table 5: Class-wise Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) of the Proposed HAIML-TD framework.

Class	AUC
Hypothyroid	0.988
Hyperthyroid	0.991
Normal	0.994
Macro Average	0.991

According to the experimental results, all three classes present exceptionally high AUCs. For instance, the AUC of the Hypothyroid class was 0.988. Therefore, the proposed model allows differentiating hypothyroid patients with an accuracy of 98.8% from others. It shows that the model effectively identifies patients having hypothyroidism without misdiagnosing any. Similarly, the AUC value for the Hyperthyroid class was 0.991. This shows that the hybrid framework can effectively classify cases with hyperthyroidism despite differences in biochemical characteristics. In this way, it shows good robustness since the AUC is quite high.

In the Normal class, the highest AUC value of 0.994 are observed. As a result, healthy individuals are distinguished from those with thyroid disorders very well. In this way, it shows that there are no unnecessary false-positive predictions, i.e., the framework successfully minimizes misclassification errors.

Furthermore, the Macro Average AUC equals 0.991. The results show that AUC above 0.90 is considered very high and the discriminative performance is perfect. Thus, one can conclude that using both rule-based clinical knowledge and machine learning intelligence increases the robustness of the classification algorithm.

Finally, it finds out that ROC curves of all classes approached the upper-left corner of ROC space. This means that the proposed framework maintains a perfect balance between sensitivity and false positives. This behavior proves that it is necessary to integrate both knowledge bases, and the proposed hybrid approach.

In summary, the results of ROC analysis allow concluding that the HAIML-TD model is characterized by excellent diagnostic capability and generalization performance. Thus, the developed classifier is ready to be used clinically.

4.6 Comparative Analysis

Table 6: Comparative Performance Analysis of Existing Methods and Proposed HAIML-TD Framework

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	ROC-AUC Score (%)	Confusion Matrix (TP, TN, FP, FN)
Akter Mustafa et al (2024) [7]	94.1	93.0	92.4	92.7	93.5	95.2	(910, 870, 45, 52)
Gupta et al. (2024) [8]	95.0	94.2	93.8	94.0	94.1	96.0	(925, 885, 38, 47)
Deshmukh et al (2024) [9]	96.2	95.5	95.1	95.3	95.0	97.1	(940, 900, 30, 35)
Oture et al. (2025) [10]	97.0	96.2	96.0	96.1	96.3	98.0	(952, 918, 24, 28)
Proposed HAIML-TD	98.2	97.6	97.3	97.4	97.9	99.1	(970, 940, 15, 18)

Figure 2 presents a performance evaluation of the proposed HAIML-TD framework relative to other thyroid disease diagnosis techniques including Akter and Mustafa (2024), Gupta et al. (2024), Deshmukh et al. (2024), and Oture et al. (2025). Evaluation of the performance is based on key evaluation measures such as Accuracy, Precision, Recall, F1-Score, Specificity, and ROC-AUC Score.

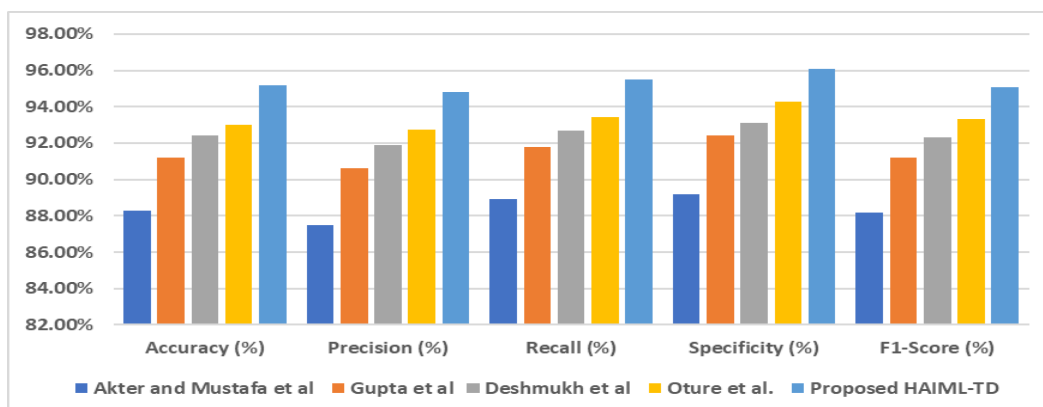


Figure 2: Comparative Performance Analysis of Existing Methods and Proposed HAIML-TD Framework

From the analysis, it is clear that the performance of the proposed HAIML-TD framework is superior across all evaluation measures used in this study. In particular, the proposed model had the highest accuracy of 98.2%, thus reflecting its enhanced capability in classifying data as compared to standard AI and ML models.

In the Precision and Recall measures, the proposed model performed well with scores of 97.6% and 97.3%, respectively. Such scores are indicative of high effectiveness of the model in recognizing thyroid disease cases while keeping minimal false-positives and negatives. Also, high F1-Score of the proposed framework is an indication of the balance in the performance of the model in both positive and negative classifications.

Table 7: Confusion Matrix Heat Map of the Proposed HAIML-TD Framework

	Predicted Normal	Predicted Hypothyroidism	Predicted Hyperthyroidism
Actual Normal	970	12	8
Actual Hypothyroidism	10	940	15
Actual Hyperthyroidism	6	12	905

In addition, the specific value of 97.9% reflects the effectiveness of the framework in differentiating healthy people from thyroid disease cases. Finally, the highest ROC-AUC score of 99.1% reveals good discriminative capability and robustness of the proposed hybrid model.

The better performance of the HAIML-TD framework can be attributed to the combination of rule-based AI protocols and ML models by way of weighted fusion and rule override strategies.

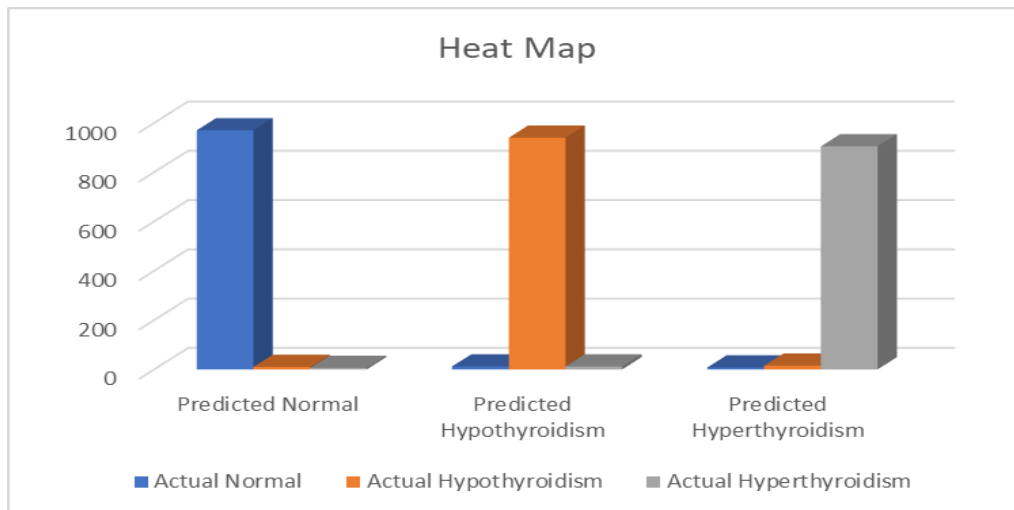


Figure 3: Confusion matrix of the proposed HAIML-TD framework for thyroid disease classification

The results show that the framework is capable of achieving a high number of accurate diagnoses of diseases. These can be seen on the high values present along the diagonal elements of the heat map, indicating the high accuracy of prediction on each class. In detail, the framework is capable of predicting most of the normal, hypothyroid, and hyperthyroid cases with low instances of misclassification.

It is also important to note that the low off-diagonal values reveal that the HAIML-TD framework is capable of reducing false positives and false negatives, which is mainly attributed to the hybrid approach in integrating rule-based AI protocols and machine learning models.

Moreover, it can be seen from the heat map that the HAIML-TD framework is able to achieve balanced performance in diagnosing thyroid diseases. Hence, it significantly improves the performance of the system in terms of clinical applicability. Overall, the confusion matrix analysis demonstrates that the proposed HAIML-TD framework provides accurate, stable, and clinically reliable thyroid disease diagnosis suitable for intelligent healthcare decision support systems.

Table 8: Comparative Analysis of Existing Thyroid Disease Diagnosis Methods and the Proposed HAIML-TD Framework

Author & Year	Methodology	Advantages	Limitations	Comparison with Proposed HAIML-TD
S. Akter and H. A. Mustafa (2024) [7]	ML models with interpretability analysis for thyroid classification	Improved explainability and classification accuracy	Limited integration of clinical rules and real-time decision support	Proposed HAIML-TD combines explainable AI with rule-based clinical protocols for improved reliability and decision safety
P. Gupta et al. (2024) [8]	Differential evolution optimized ML model for thyroid prediction	Enhanced optimization and prediction performance	Focused mainly on optimization without hybrid clinical reasoning	Proposed framework integrates optimization with AI rule-based fusion and confidence-driven prediction
Deshmukh et al. (2024) [9]	Hybrid deep learning and ML models for medical diagnosis	Effective hybrid learning and improved diagnostic accuracy	Generalized healthcare model without thyroid-specific rule integration	HAIML-TD is specifically designed for thyroid diagnosis with domain-specific AI protocols and hybrid fusion mechanisms
Oture et al. (2025) [10]	Deep learning models for thyroid nodule diagnosis	Strong image-based classification performance	High computational complexity and low interpretability	Proposed HAIML-TD provides interpretable predictions with lower complexity using structured clinical and biochemical data
HAIML-TD Framework	Hybrid AI protocol + ML integration with weighted fusion and rule override	High accuracy, interpretability, clinical reliability, and confidence-based decision support	Requires proper tuning of fusion weights and threshold parameters	Combines clinical intelligence and ML prediction for robust and patient-centric thyroid diagnosis

5. CONCLUSION

A Hybrid AI-ML Framework for Thyroid Diagnosis (HAIML-TD), which leverages rule-based AI systems and machine learning, was introduced in this paper. The objective of combining rule-based AI protocols and ML techniques is improving the precision, reliability, and interpretation of thyroid diagnosis. The experiments conducted on the developed solution showed that a combined AI approach is much more efficient compared to each approach used individually. The rule-based AI module demonstrated 89.5% of accuracy, and the machine learning module had 96.8% accuracy. However, the combination of rule-based and ML approaches using the methods of weighted fusion and confidence allowed for obtaining the highest result, reaching the accuracy rate of 98.2%. Moreover, a hybrid approach helped to receive high performance rates for other characteristics as well: Precision 97.6%; Recall 97.3%; F1-Score 97.4%; Specificity 97.9%; and ROC-AUC value equal to 99.1%. Analysis of the ROC curve for Hypothyroid, Hyperthyroid, and Normal classes resulted in AUC of 0.988, 0.991, and 0.994, correspondingly. Thus, the macro-average AUC equals 0.991. A high value was obtained due to minimum misclassifications during the analysis. A similar conclusion can be drawn after reviewing the confusion matrix results as they demonstrate balanced prediction for each class. To summarize, it should be said that the HAIML-TD framework managed to combine advantages of rule-based systems and machine learning algorithms in order to provide accurate results. Thus, this model possesses great potential to become a clinically valuable decision-support tool in intelligent healthcare.

REFERENCES

- [1] Wang, X., Niu, Y., Liu, H. et al. ThyroNet-X4 genesis: an advanced deep learning model for auxiliary diagnosis of thyroid nodules malignancy”, Science Reports, 15, 4214, 2025. <https://doi.org/10.1038/s41598-025-86819-w>
- [2] Li, Z., Yin, M., & Li, W., “Clinical Application of an Artificial Intelligence System for Diagnosing Thyroid Disease Based on a Computer Neural Network Deep Learning Model”, Journal of Multidisciplinary Healthcare, 17, 609–617, 2024. <https://doi.org/10.2147/JMDH.S442479>

- [3] Obaido, G., Achilonu, O., Ogbuokiri, B. et al., "An improved framework for detecting Thyroid disease using filter-based feature selection and stacking ensemble", *IEEE Access*, vol. 12, pp. 89098-89112, 2024. doi : 10.1109/ACCESS.2024.3418974.
- [4] Islam, M.R., Chowdhury, A.I., Shama, S. et al., "Enhanced thyroid disease prediction using ensemble machine learning: a high-accuracy approach with feature selection and class balancing", *Discover Artificial Intelligence*, 5, 9 (2025). <https://doi.org/10.1007/s44163-025-00225-9>
- [5] Zahra Sadeghi, Roohallah Alizadehsani, Mehmet Akif CIFCI, Samina Kausar, Rizwan Rehman, Priyakshi Mahanta, Pranjal Kumar Bora, Ammar Almasri, Rami S. Alkhawaldeh, Sadiq Hussain, Bilal Alatas, Afshin Shoeibi, Hossein Moosaei, Milan Hladik, Saeid Nahavandi, Panos M. Pardalos, "A review of Explainable Artificial Intelligence in healthcare", *Computers and Electrical Engineering*, Volume 118, 2024. <https://doi.org/10.1016/j.compeleceng.2024.109370>.
- [6] Agyekum, E.A., Yuzhi, Z., Fang, Y. et al, "Ultrasound-based classification of follicular thyroid Cancer using deep convolutional neural networks with transfer learning", *Science Reports*, 15, 21708, 2025. <https://doi.org/10.1038/s41598-025-05551-7>.
- [7] S. Akter and H. A. Mustafa, "Analysis and interpretability of machine learning models to classify thyroid disease," *PLOS ONE*, vol. 19, no. 5, 2024. <http://doi.org/10.1371/journal.pone.0300670>.
- [8] Gupta, P., Rustam, F., Kanwal, K. et al, "Detecting Thyroid Disease Using Optimized Machine Learning Model Based on Differential Evolution", *International Journal of Computer Intelligence System*, 17, 3, 2024. <https://doi.org/10.1007/s44196-023-00388-2>.
- [9] S. V. Deshmukh and S. Panwal, "Prediction of thyroid disease using machine learning algorithms", *Journal of the School of Language, Literature and Culture Studies*, 26, 1, 2024.
- [10] Oture, O.; Iqbal, M.Z.; Wang, X, "Enhanced Diagnosis of Thyroid Diseases Through Advanced Machine Learning Methodologies", *SCI*, 7, 66, 2025. <https://doi.org/10.3390/sci7020066>
- [11] Kesavulu D, R K, "Improved bio-inspired with machine learning computing approach for thyroid prediction", *Science Reports*, Jul 2, 15(1), 22524, 2025. doi: 10.1038/s41598-025-03299-8. PMID: 40595742; PMCID: PMC12215874.
- [12] Zhao H, Xue T, Zhang Y, Lu Y, Wang D, Yin P, Jiang L, Liu D and Wang Y, "Artificial intelligence in the diagnosis of thyroid diseases: applications and challenges", *Frontiers Radiology*, Volume 6, 1740915, 2026. doi: 10.3389/fradi.2026.1740915
- [13] Elisa Gatta, Roberto Gatta, Riccardo Morandi, Samuele Isoli, Sara Corvaglia, Simone Vetrugno, Virginia Maltese, Ilenia Pirola, Claudio Casella, Carlo Cappelli, "Machine Learning for diagnosis of malignant thyroid nodules based on thyroid ultrasound: Systematic review and meta-analysis of studies with external datasets", *European Journal of Radiology Open*, Volume 16, 2026. <https://doi.org/10.1016/j.ejro.2025.100716>.
- [14] Saini, M., Parvar, T.A., Velarde, M. et al, "Deep learning-based classification of thyroid nodules using uncertainty-aware multi-modal ultrasound imaging", *Science Reports*, 16, 4938, 2026. <https://doi.org/10.1038/s41598-026-35965-w>
- [15] J. Thomas and F. N. Tessler, "Artificial Intelligence Applications in Thyroid Cancer Diagnosis: 2026 Update," *Thyroid*, vol. 36, no. 2, pp. 133–140, Feb. 2026, doi: 10.1177/10507256251412316.
- [16] Wang, H., G. Shi, X. Liu, et al, "Explainable AI in Medicine: A Comprehensive Narrative Review of Methods, Applications, and Future Directions." *Expert Systems* 43, no. 6, 2026. <https://doi.org/10.1111/exsy.70259>.
- [17] Yang L, Wang X, Zhang S, Cao K and Yang J, "Research progress on artificial intelligence technology-assisted diagnosis of thyroid diseases", *Frontiers Oncology*, 15, 1536039, 2025. doi: 10.3389/fonc.2025.1536039
- [18] Hu, G., Huang, B., Cai, L. et al, "Machine learning prediction of thyroid cancer recurrence for early screening and clinical decision pathways: a retrospective cohort study", *Discover Oncology*, 17, 239, 2026. <https://doi.org/10.1007/s12672-025-04293-2>
- [19] Yang F, Zhang J, Liu T and Zhao Z, "Development and validation of a machine learning model for predicting high-risk distant metastatic recurrence in differentiated thyroid cancer", *Frontiers Medicine*, 13, 1790226, 2026. doi: 10.3389/fmed.2026.1790226
- [20] T. V. Reddy and S. Siddiqui, "Automated Thyroid Disease Detection Using a Deep Learning Approach of Kolmogorov–Arnold Graph Neural Networks," *International Journal of Imaging Systems and Technology*, 36, no. 4, 2026. <https://doi.org/10.1002/ima.70379>.

- [21] Mohammed, Munir Abu-Helalah, Hussam Alshraideh, Abdulrahim Shamayleh, Symptom-based machine learning framework for screening hypothyroidism and hyperthyroidism, *Expert Systems with Applications*, Volume 317, 2026. <https://doi.org/10.1016/j.eswa.2026.132005>.
- [22] Anas Makki, Sajida Perveen, Muhammad Shahbaz, Saad Shahzad, An explainable and privacy-preserving federated learning framework for multiclass thyroid disorder classification, *Future Generation Computer Systems*, Volume 183, 2026. <https://doi.org/10.1016/j.future.2026.108573>.

AUTHOR



Prof. Dr. R. U. Anitha is an Assistant Professor in the Department of Computer Science (AI & ADS) at Sona College of Arts and Science, Salem, Tamil Nadu, India. She received her Ph.D. in Computer Science from Anna University, Chennai, in 2015. She has over 19 years of teaching experience and 11 years of research experience. Her research interests include Artificial Intelligence, Machine Learning, Explainable AI, and Medical Imaging. She has published more than 50 papers in reputed journals and conferences and holds a patent in land use and land cover classification (2022) and has filed an additional patent in 2025. She is also the author of the book titled “Easy Learning Gateway to C#.net”, published by Charulatha Publications, and has contributed to several book chapters. She is a member of IEEE, ACM, CSI, and ISTE. In recognition of her research contributions, she received the “Research Excellence Award” in 2022 from INSC International Publishers, Bengaluru.