Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 3, Issue 9, 2025

PRAGMATIC TECHNIQUES FOR THYROID DISEASE CLASSIFICATION USING MACHINE LEARNING

Hina Shahid[1], Zahida Yaseen[*2], Muhammad Ahmad Hanif[3], Shoaib Saleem[4], Zahid Hasan[5], Gulzar Ahmad[6]

[1]Department of Computer Sciences, University of Gujrat, 50700, Pakistan
*[*2]Center for Research and Development (CRD), Minhaj University, Lahore 54000, Pakistan*
[3,4,6]Department of Computer Science, Minhaj University, Lahore 54000, Pakistan
[5]Department of Computer Science, National College of Business & Economics, Lahore 54000, Pakistan.

DOI:https://doi.org/10.5281/zenodo.17096736

Keywords
Machine Learning; Classification; Thyroid disease; XGBoost; Pragmatic techniques; Thyroid disease

Article History
Received: 18 June 2025
Accepted: 28 August 2025
Published: 11 September 2025

Copyright @Author

Corresponding Author: *
Zahida Yaseen

Abstract
The thyroid gland, a vital regulator of the body's metabolism, releases hormones into the bloodstream to control bodily functions. Hypothyroidism and hyperthyroidism are two hormone disorders associated with the thyroid. Factors such as iodine deficiency, autoimmune conditions, and inflammation can contribute to thyroid issues. Diagnosis involves a blood test, but data noise is common. Implementing data cleaning techniques simplifies analytics, allowing for a clearer assessment of the patient's risk of developing thyroid disease. In this study, machine learning techniques were employed for the classification of thyroid diseases using a dataset from Kaggle. The thyroid's crucial role in essential functions underscores the significance of early disease diagnosis. Multiple machine learning algorithms were implemented, and their performance was assessed through confusion matrices. Pragmatic techniques were applied for data optimization, contributing to the development of robust models. A comprehensive analysis compared various metrics, revealing that the XGBoost Classifier (XGB) demonstrated the highest effectiveness with 99.46% accuracy. The study highlights the importance of tailoring techniques to dataset characteristics, available resources, and expertise. Addi-tionally, comparisons with established models showcase notable improvements in results and procedures

1. INTRODUCTION

Health is the cornerstone of a progressive society and a fundamental human right. When individuals enjoy good health, they can actively contribute and participate in their societal roles, enabling them to reach their full potential. Healthy individuals can learn and work productively, making health a valuable asset for any competitive and rapidly developing society. One health condition that poses challenges to an individu-al's daily activities is thyroid dysfunction. The thyroid, a small gland located in the neck, is often overlooked despite its crucial role in maintaining overall health and well-being. Comprising lobes situated at the front of the neck, the thyroid serves as the central part of the gland, while the parathyroid consists of two lobes on either side. The thyroid gland regulates bodily growth and various physiological processes by releasing thyroid-stimulating hormones into the bloodstream. These hormones interact with dif-

ferent organs and regions of the body, influencing factors such as heart rate, body temperature, metabolism, and cognitive functions [1]. Thyroid illness can affect individuals of all ages, including men, women, and children. In the United States alone, approxi-mately 20 million people suffer from thyroid dysfunction, indicating its relatively high prevalence [2]. Furthermore, research has shown that women are more likely to expe-rience thyroid issues compared to men. The thyroid gland produces two primary hor-mones, Triiodothyronine (T3) and thyroxine (T4), which are synthesized from dietary iodine. These hormones play a crucial role in regulating the body's metabolism. The hypothalamus and pituitary glands in the brain work together to maintain the balance of T3 and T4 levels. The hypothalamus produces TSH Releasing Hormone (TRH), which signals the pituitary gland to stimulate the thyroid to produce the required amount of T3 and T4 by either increasing or decreasing the release of thyroid-stimulating hor-mone (TSH).

In Pakistan, clinical and subclinical hypothyroidism affects approximately 4.1% of adults and 5.4% of children, with both conditions being more prevalent in females than males [3]. Early detection, diagnosis, and treatment of thyroid disease are crucial for limiting disease progression. Timely identification and differential diagnosis enhance the chances of effective therapy for various abnormalities. Despite numerous efforts, clinical diagnosis is often perceived as challenging [4]. Information technology plays a vital role in disease detection and diagnosis through the utilization of analytical tools. In comparison to other industries, the healthcare sector has been slower to adopt in-formation technology [5]. Among the available analytical technologies, Artificial Intel-ligence (AI) is hailed as a powerful and beneficial tool for humanity [6]. In healthcare, AI involves identifying patterns within small groups or clusters of individuals and forecasting the clinical course for the rest of the population. Ingeniously, epidemics of infectious diseases have been predicted using diverse datasets. However, the unpre-dictable nature of contagious diseases and the multitude of variables influencing infec-tion and transmission present challenges. AI is a branch of computing that enables machines to perform cognitive tasks akin to humans. In AI methodologies, machines are designed to analyze, interpret, and solve problems. Machine Learning (ML) is a popular subset of AI, wherein computers learn to behave or respond based on specific outcomes and subsequently make decisions and predictions using acquired knowledge and experiences [7]. For ML, a computer program is assigned tasks, and as the pro-gram's measured performance improves with increasing experience, it is inferred that the machine has learned. Consequently, the machine begins making decisions and forecasts based on factual information. For instance, computer software that learns to identify or predict cancer based on relevant medical reports of patients gradually im-proves its performance by analyzing medical records from a larger population. Per-formance indicators in such cases could include the number of accurate cancer predic-tions compared

2. Problem Statement

Thyroid disease classification is a complex task that requires accurate and timely iden-tification of different thyroid disorders. Traditional diagnostic methods often rely on manual interpretation and subjective assessments, leading to potential errors and de-lays in diagnosis. To overcome these challenges and improve the efficiency and accu-racy of thyroid disease classification, there is a need for pragmatic techniques that lev-erage machine learning (ML) algorithms. The main problem addressed in this study is the lack of pragmatic techniques and comprehensive evaluations of ML algorithms de-signed explicitly for thyroid disease classification. Developing and evaluating such techniques would provide healthcare practitioners with reliable and efficient tools to aid in the accurate diagnosis of thyroid diseases, leading to improved patient care and management. Additionally, identifying the most suitable ML algorithms and their op-timal configuration will enable the integration of ML-based systems into clinical prac-tice, ensuring their effective utilization and acceptance by healthcare professionals.

2. Research Motivation

The motivation for this study has taken this research to the following question;

- By using ML algorithms, the accuracy and efficiency in thyroid disease classification can be enhanced, ensuring that healthcare professionals perform reliable decisions and prescribe suitable treatment.
- The design of these practical methods could detect thyroid diseases earlier which will result in early management and better patient's outcome. There is a possibility that the use of A BL and ML techniques could anal-yse large data sets and detect patterns that indicate thyroid pathological con-ditions in the early stages.
- ML-based (machine learning) systems can help in detection of disease and help doctors by automated classification of thyroid diseases, thus making them focus on more critical parts for the time-being of patient care. Pragmatic approaches may be able to provide useful and user-friendly interfaces, and help ML to become part of clinical work-flows.
-

4. Significance of our Study

Overall, the significance of our study lies in its potential to advance the field of thyroid disease classification, improve diagnostic accuracy, facilitate timely interventions, op-timize healthcare resource utilization, enhance decision support for healthcare profes-sionals, and bridge the gap between research and clinical practice. By addressing these aspects, our study aims to contribute to improved patient care and outcomes in the context of thyroid diseases.

5. Research Objectives

The following objectives are defined for this research;

a) Gather a comprehensive dataset of thyroid disease cases, including relevant clin-ical and diagnostic information. Preprocess the data to ensure its quality, consistency, and compatibility with machine learning algorithms.

b) Assess the performance and effectiveness of the developed pragmatic techniques through rigorous evaluation and validation processes. Utilize additional datasets or conduct experiments with clinical data to ensure the generalizability and reliability of the proposed techniques.

c) Develop pragmatic techniques that address thyroid disease classification's specific challenges and limitations in real-world clinical settings. Consider interpretability, simplicity, and ease of integration into existing healthcare systems.

6. Literature Review

Chaubey, Gyanendra, et al. conducted a study utilizing multiple ML classifiers, such as AdaBoost, Gradient Boosting, XGBC, CatBoost, NB, KNN, MLP, DT, ExtraTree, and Random Forest (RF). The dataset consisted of 2,211 instances from 247 patients and was collected from the Naples hospital. The findings revealed that the EXTC model outperformed other methodologies, achieving an F-score of 84% [9]. In another study by authors in [13], multiple ML algorithms including SVM, KNN, and DTs were ap-plied to a dataset obtained from the UCI ML repository to forecast the risk of thyroid cancer development in patients. According to their results, SVM demonstrated the best performance with an accuracy rate of 99.63% and showed promise in detecting thyroid disease. Borzouei et al. focused on the most prevalent thyroid disorders, hyperthyroid-ism and hypothyroidism, in their study [14]. They employed neural network classifica-tion and multinomial logistic regression models, using demographic and hormonal in-formation as input. The most accurate predictive models were found to be those using logistic regression [15]. In [12], authors analyzed the performance of various ML classifiers using two datasets, "Conventional US" and "Conventional US and RTE," based on clinical and pathological data from 2,064 thy-roid nodules. The goal was to distinguish between benign and malignant nodules. The RF algorithm performed the best and demonstrated high accuracy not only based sole-ly on US features but also on US and RTE features. In their work [16], Temurtas evalu-ated and compared linear

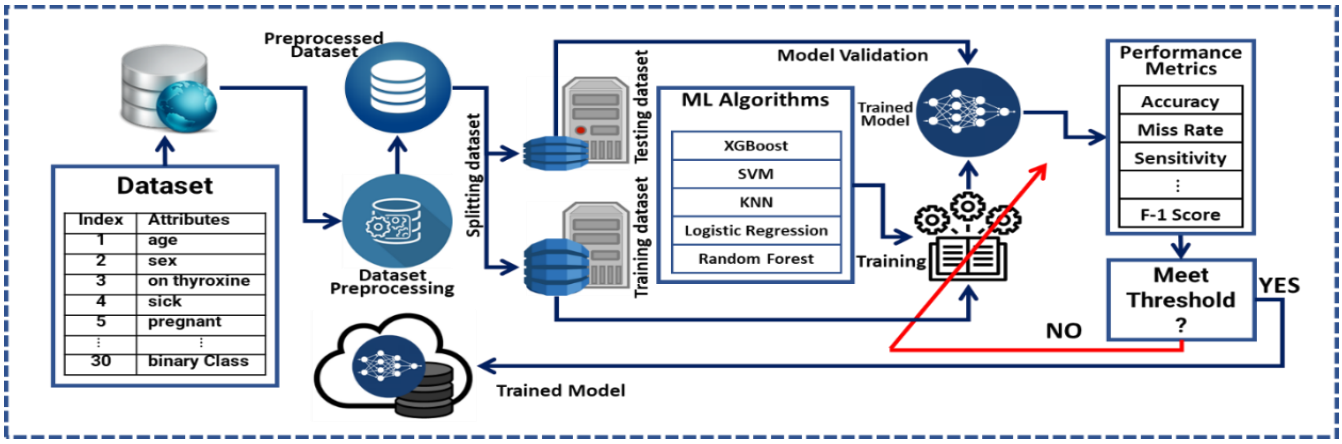
and nonlinear ML algorithms to assess the risk of thyroid nodule malignancy. The study utilized data from a medical database (2011-2016) for training and validation. Linear ML algorithms such as Ridge, Lasso-penalty, and Elas-tic Net (EN) were analyzed, along with nonlinear algorithms such as RF, ker-nel-Support Vector Machines (k-SVM), Neural Network (N net), KNN, and NB. The RF algorithm outperformed other ML algorithms in the training cohort (0.989 (95% CI: 0.975–1.000)), and the RF and SVM algorithms outperformed different algorithms in the validation set (95%). ML algorithms also showed higher AUCs compared to expe-rienced radiologists working independently (AUC, 0.830, 95% CI: 0.811-0.849).

7 Proposed Methodology

This portion exemplifies the functioning of proposed model by describing the dataset and its preprocessing, ML algorithms used for training the model, Model validation by using Performance Metrics, language and Application Module used for this research.

M.A.A.R. Asif et al. applied various ML algorithms, including KNN, SVM, XGB, Ada-Boost (AdB), Gaussian Process Classifier (GPC), Gradient Boosting Classifier (GBC), and Multiplier Perceptron Classifier (MLPC), to the UCI thyroid illness dataset. The Multilayer Perceptron (MLPC) classifier achieved the highest accuracy of 99.70% [17]. Abbad Ur Rehman et al. employed ML algorithms such as NB, LR, KNN, SVM, DT, and various performance metrics on data collected from a Teaching Hospital in Dear Gha-zi Khan, Pakistan. The best accuracy achieved was 100% for NB and LR [18]. Duggal et al. in [19] utilized SVM and Recursive Feature Elimination (RFE) models to detect thy-roid disease, achieving an accuracy of 92.92%

Figure 1: Conceptual description of the proposed solution.



The data set utilized for the experimental purpose is obtained from the Kaggle repository website (web source https://www.kaggle.com/code/vuppalaadithyasai data was imbalanced, as shown in Figure 1 of this paper. Therefore, preprocessing of the dataset is required. Dataset has been preprocessed by:

ram/98-test-accuracy-thyroid-prediction/data). A dataset comprises 30 attributes with 3220 rows, having numerical and missing values. In addition,

- Handling missing values

- Handling categorical data
- Removing duplicate data
- Normalization

Further, split this preprocessed dataset into two datasets, 'Training Dataset' and 'Testing Dataset,' in the 90% and 10% ratio, respectively. Detail of the attributes from the dataset mentioned above is listed in tabular form in Tab.1.

System-level monitoring layer: Based on the different configurations, the node adjustment, and the status of VMs are analyzed by the monitoring system with the ability to ensure the stable operation of the entire architecture. In the event of node failure the load is shared by making use of the nearby virtual machines to commit an instance for the process recovery to recover the system once more. Inbalanced dataset inbalanced dataset is a dataset that has a oor balance between classes and i.e. one class is represented by a large number of examples whereas others are rbitrary few. In other words, the number of instances belonging to different classes in the dataset is imbalanced or unequal as shown in Fig. 2.

Table:1 Attributes details

Attribute	Value Type	Explanation
Age	Integer	-
Sex	M (male), F (female)	-
on thyroxine	t (true), f (false)	-
query on thyroxine	t (true), f (false)	-
on antithyroid medication	t (true), f (false)	-
Sick	t (true), f (false)	-
Pregnant	t (true), f (false)	-
thyroid surgery	t (true), f (false)	-
I131 treatment	t (true), f (false)	-
query hypothyroid	t (true), f (false)	-
query hyperthyroid	t (true), f (false)	-
Lithium	t (true), f (false)	-
Goiter	t (true), f (false)	-
Tumor	t (true), f (false)	-
Hypopituitary	t (true), f (false)	-
Psych	t (true), f (false)	-
TSH measured	t (true), f (false)	-
TSH	Real	Thyroid Stimulating Hormone
T3 measured	t (true), f (false)	-
T3	Real	Triiodothyronine
TT4 measured	t (true), f (false)	-
TT4	Real	Total T4 (Thyroxine)
T4U measured	t (true), f (false)	-
T4U	Real	T4 uptake
FTI measured	t (true), f (false)	-
FTI	Real	Free T4 Index
TBG measured	t (true), f (false)	-
TBG	Real	thyroxine binding globulin
referral source	SVHC, SVI, STMW, SVHD, other	-
binaryClass	P (positive), N (negative)	-

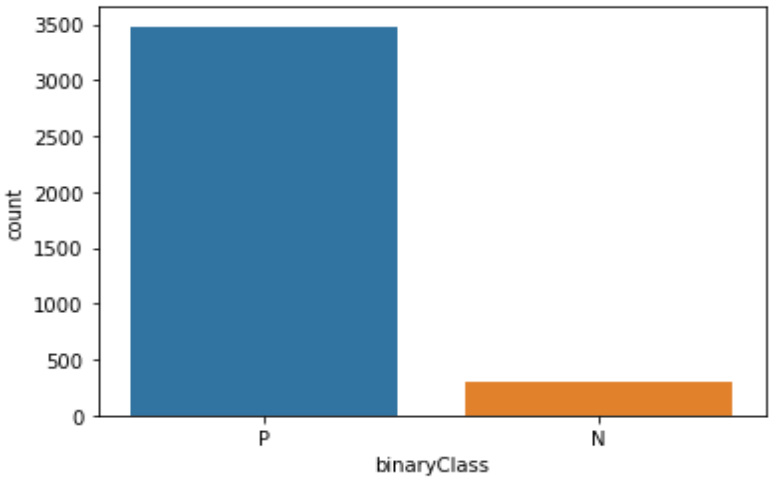


Figure 2: Imbalanced Dataset

In this proposed Model, the multiple ML Algorithms including XGB, SVM, KNN, RF, LR, DT, NB are implemented on training dataset (90%) to train the Model and observe the performance of each ML Algorithm. After data preprocessing and training of the model, trained model is validated with the remaining 10% test dataset.

Execution of these ML algorithms are executed in python coding. Detail of these ML algorithm are given below:

7.1 XGBoost (XGB)

XGB has gradient boosting as its main component, based on DTs as a classifier. In addition, it uses a multi-threaded technique that efficiently uses the computer's CPU core, resulting in increased speed and performance. It has been employed because it is quick, effective, and scalable [20], [21].

The results of this paper show that the XGB algorithm performs better than other ML Algorithms.

7.2 Support Vector Machine (SVM)

SVM is famous supervised ML algorithms, which is employed for both classification and regression problems, largely used for classification analysis. Formally speaking, an immaculate hyperplane describes SVM. Support vector is a type of data set that supports a hyperplane. SVM attempts to specify the best hyperplane with the shortest distance to the training set. Finally, we get the hyperplane equation $\omega.\alpha + b = 0$. The hypothesis function for SVM classifier is fol: [22]

$$\text{Hypothesis } (\alpha_i) = \begin{cases} +1 \text{ if } \omega.\alpha \geq 0 \\ -1 \text{ if } \omega.\alpha < 0 \end{cases}$$

7.3 k- Nearest Neighbor (KNN)

KNN mostly used for classification problems. It keeps all available cases and classifies new cases on a similar basis. It is suitable where dataset is small, data is noise free and labeled.

KNN is mainly used for classification problems. It stores all of the cases that are accessible and classifies new cases based on similarity. It is suitable where the dataset is small; data is noise-free and labeled. KNN categorized objects based on the class of their closest neighbors. The method is often known as KNN Classification, where the class is determined by the k nearest neighbors [23].

It employed the Euclidean distance formula to compute the distance between two endpoints in the plane with coordinates (x, y) and (α, β) mentioned below:

$$dist(d) = \sqrt{(x_i - \alpha)^2 + (y_i - \beta)^2}$$

7.4 Decision Tree (DT)

A visual representation of a set of options and their results in the shape of a tree-like structure is called a decision tree (DT). The DT comprises nodes and edges, where nodes stand in for events or decisions and edges for the circumstances or rules governing those decisions. Each branch in the tree represents a potential value that the node

might take, and each node represents a group of qualities that need to be categorized [24].

7.5 Naive Bayes (NB)

Naive Bayes is a classification method that utilizes the Bayes Theorem while assuming that the predictors are independent. In other words, it believes that the presence of any other part does not influence the existence of a unique feature in a class. This technique is often used for text classification applications and primarily focuses on clustering and classification based on the conditional probability of an event occurring.[24].

7.6 Logistic Regression (LR)

It is a supervised ML algorithm designed to estimate the possibility of a particular result or observation for a dependent variable based on input/ independent variable by fitting input to a logistic function. The Logistic regression technique uses the famous sigmoid function and provides a chance of incidence between 0 and 1, which is described as follows: [25]–[27]

$$f(a) = \frac{1}{1 + e^{-a}}$$

7.7 Linear Discriminant Analysis (LDA)

A supervised classification technique to reduce the number of areas is linear discriminant analysis (LDA). In ML and pattern recognition applications, it operated as a preprocessing step. LDA conserves class-discriminatory information while projecting characteristics onto a lower dimensional space.

7.4 Model Validation Using Performances Metrics

Multiple performance metrics are applied to training and test datasets, divided at a ratio of 90% and 10%, respectively, to measure the accuracy and performance of the proposed **PT2DC** model. These performance metrics are named Accuracy, True Positive Rate (TPR), False Negative Rate (FNR), True Negative Rate (TNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), False Positive Rate (FPR), False Discovery Rate (FDR), and F1-Score. The Confusion Matrix in table 2 presents the findings of training and test datasets.

The confusion matrix demonstrates the degree of accuracy of the model. The accuracy of the proposed model **PT2C** has been calculated by using confusion matrix's accuracy formula is shown in eq. 1.

$$Accuracy = \frac{(\xi_\rho + \xi_\eta)}{N} \quad 1$$

(Where ξ_ρ true positive, ξ_η true negative and N total number of dataset)

The suggested model's error rate, miss rate, or false negative rate (FNR) is calculated by erroneously identifying in the confusion matrix is given in eq. 2.

$$FNR \text{ / } Miss\ rate = \frac{(\zeta_\rho + \zeta_\eta)}{N} \quad 2$$

(Where ζ_ρ False positive, ζ_η False negative)

eq. 3 calculates the sensitivity, recall, or true positive rate (TPR).

$$TPR \text{ / } Sensitivity = \frac{\xi_\rho}{(\xi_\rho + \zeta_\eta)} \quad 3$$

Other performance metric is Specificity or True Negative Rate (TNR) is determined by eq. 4.

$$TNR \text{ / } Specificity = \frac{\xi_\eta}{(\xi_\eta + \zeta_\rho)} \quad 4$$

eq. 5 calculates the proposed model's precision or positive predictive value

$$Precision = \frac{\xi_\rho}{(\xi_\rho + \zeta_\rho)} \quad 5$$

Negative Predictive Value (NPV) of the proposed PT2DC model is determined by eq. 6.

$$NPV = \frac{\xi_\eta}{(\xi_\eta + \zeta_\eta)} \quad 6$$

eq. 7 measures False Positive Rate (FPR) of Fall-Out of the proposed PT2DC model

$$FPR = \frac{\zeta_\rho}{(\xi_\eta + \zeta_\rho)} \quad 7$$

False Discovery Rate (FDR) of the proposed PT2DC model is determined by eq. 8.

$$FDR = \frac{\zeta_\rho}{(\zeta_\rho + \xi_\rho)} \quad 8$$

The harmonic mean of precision and recall calculates the F1 score as shown in eq. 9. F1-Score is a crucial indicator for assessing the proposed model.

$$F1 - Score = \frac{2 * TPR * PPV}{TPR + PPV} \quad 9$$

In Our proposed PT2DC model, ML algorithms are trained to avoid bias. Therefore, they are free from overfitting as the dataset has been used after preprocessing and optimizing by removing duplicate data, normalizing data, and handling missing values and categorical data.

A confusion matrix is a matrix that shows the performance of a classification model by displaying the number of correct and incorrect predictions made by the model as shown in Table 2 and Table 3.

Moreover, there is no possibility of data leakage as testing data is used to validate the model/ ML algorithms. Thus, Performance in terms of accuracy is best given by XGB of 99.46% on testing data. However, the performance matrices for DT (98.39%), KNN (96.77%), and LDC (94.35%) demonstrated promising accuracy result. A complete assessment of algorithms in relations of performance metrics and the Performance evaluation of the proposed model using test data with different statistical measures like accuracy, precision, specificity, sensitivity, NPV, FPR, FNR, FDR, and F1 score are displayed in Fig. 3 and Tab. 4.

IEER
Institute for Excellence in Education & Research

Table 2: Confusion matrix XGB Classifier

Proposed Model PT2DC (10% test data for validation)				
Total number of samples= 372		Predicted Output (O_t, O_f)		
Input	Expected Output (E_t, E_f)	O_t (True)	O_f (False)	Total
	E_t (True)	27	2	29
	E_f (False)	0	343	343
Total		27	345	372

Table 3: Confusion Matrix of DT

Proposed Model PT2DC (taking 10% test data used for validation)	
Total number of samples= 372	Predicted Output (O_t, O_f)

Input	Expected Output (E_t, E_f)	O_t (True)	O_f (False)	Total
	E_t (True)	26	3	29
	E_f (False)	3	340	343
Total		27	345	372

8. Performance Evaluation of the System

When to Compare Algorithms Quantitative values are powerful, but there are also many factors which should be considered to help determine when one algorithm is better than another. Following are some general-purpose performance criteria used for the comparison of algorithms:

Accuracy: It quantifies the goodness of the algorithm in predicting the correct class. Commonly, it is expressed as percent number of correctly classified items / total number of items. Larger values of accuracy represent better performance however, it may not always be a good metric to choose particularly in case of skewed classes.

Precision and recall are standard in binary classfication. Precision represents the ratio of true positives to all positive predictions, and recall measures the ratio of true positive predictions to all actual positive cases. When the trade-off of false positive and false negative is important, then these two criteria are useful.

F1 scores merge precision and recall to a single score and give a balanced measure of the performance of the model. It takes into account both the proportion of correctly-identified positive instances (precision) and the ability to identify all positive instances (recall). The F1 score is the harmonic mean of the precision and the recall, and it gives a single value that can be used to summarize the performance of a model as shown in Figure 4.

In addition to predictive performance, it is essential to consider the computational complexity of an algorithm. This includes factors such as time complexity (how the algorithm scales with input size) and space complexity (how much memory or storage the algorithm requires). Algorithms with lower computational complexity may be more desirable, especially when dealing with large datasets or real-time applications.Scalability refers to how well an algorithm performs as the dataset size increases. It is crucial to evaluate whether an algorithm can handle large volumes of data efficiently without compromising performance or requiring excessive computational resources. Robustness measures how well an algorithm performs under different conditions, such as noisy or missing data, outliers, or changes in data distribution. A robust algorithm is less sensitive to such variations and maintains consistent performance.

IEER
Institute of Excellence in Education & Research

When comparing algorithms, it is important to select performance metrics that align with the specific problem and objectives of the analysis. Different metrics provide different insights, and it is often beneficial to consider multiple metrics to comprehensively understand an algorithm's performance characteristics. Additionally, it is important to conduct statistical significance tests, such as cross-validation or hypothesis testing, to ensure that observed differences in performance are statistically meaningful.

Figure 3: Comparison in terms of performance metrics

Table 4: Performance evaluation

ML Algorithm	Accuracy	FNR Miss Rate	TPR Sensitivity	TNR Specificity	PPV Precision	NPV	FPR	FDR	F1-Score
XGB	99.46%	0.54%	96.55%	100%	99.71%	99.42%	0	0	99.73%
DT									
KN									
LD									
LR									
SV									
NB									

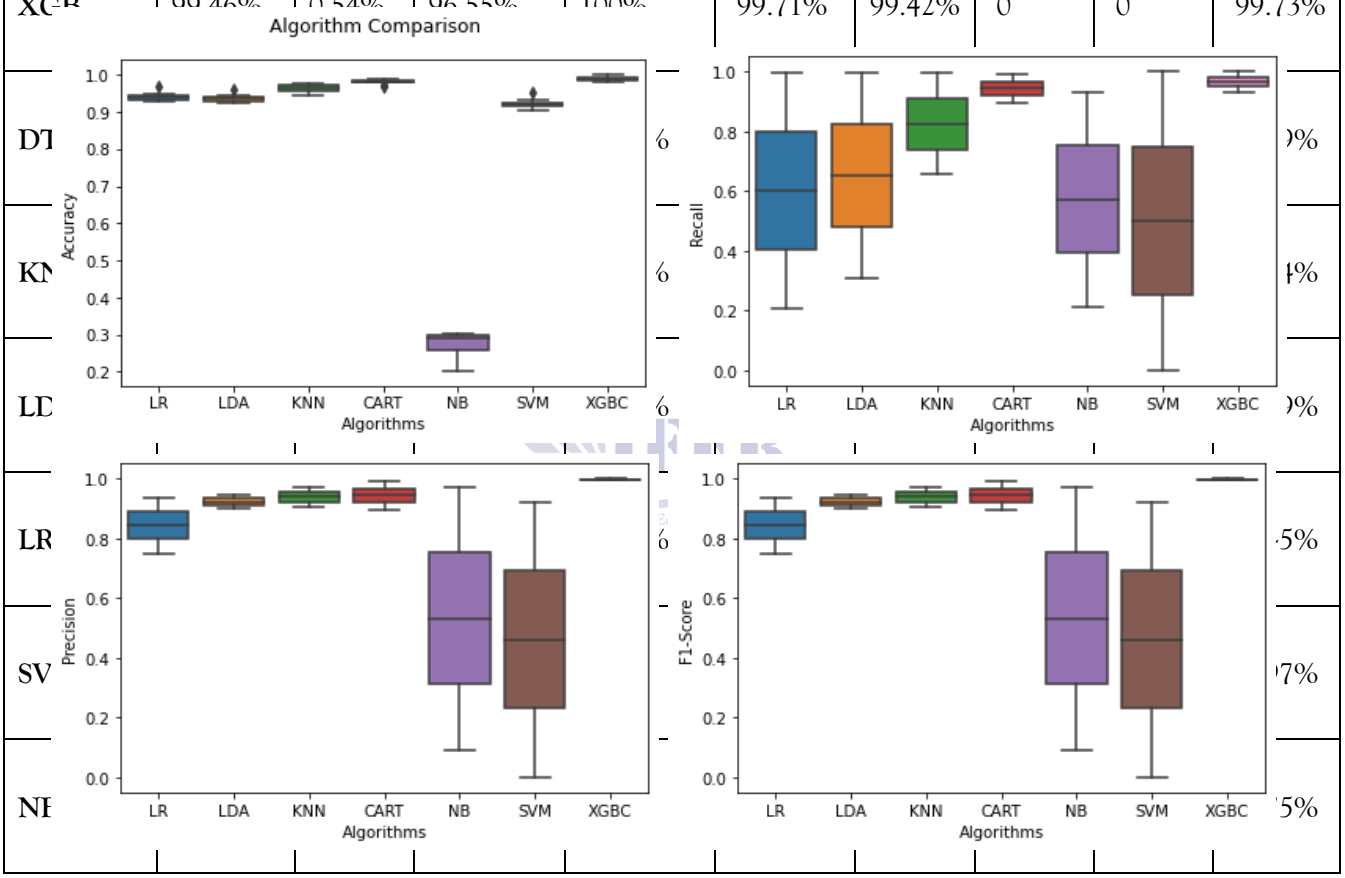


TABLE 5: Comparison with other related model in the literature

Literature	Validation/ Testing		
	Methodology	Accuracy (%)	Dataset
Alyas, Tahir, et al. (2022)	KNN	59	**UCI repository dataset thyroid**

[15]	ANN NB RF	94 93 94.8	**disease**
Asif et al. (2020) [17]	KNN SVM XGB	93.34 96.00 96.20	**UCI repository dataset thyroid disease**
Chaubey et al. (2021) [9]	KNN DT	96.87 87.5	**UCI repository dataset thyroid disease**
Proposed model (PT2DC)	XGB DT KNN LDA LR SVM	99.46 98.39 96.77 94.35 93.28 92.2	**Kaggle repository dataset thyroid disease**

9 Conclusions

Optimizing thyroid disease classification using pragmatic machine-learning approaches involves implementing practical and effective techniques to enhance the accuracy and efficiency of the classification model. A vital hormone gland, the thyroid gland, is essential for the human body's growth, development, and metabolism. The thyroid gland develops more hormones when the body entails additional energy to maintain its energy level. Thyroid dysfunction is an ailment that makes it challenging to go about their daily lives. Therefore, early disease detection using practical tools is helpful in this technological era. In this paper, we proposed an effective PT2DC model for predicting thyroid disease using different ML algorithms, with the best result given by XGB and an accuracy of 99.46%. This PT2DC model as an application is helpful in healthcare departments and for doctors to early diagnose thyroid disease more effectively and without delay. Therefore, our proposed PT2DC model is more practical in achieving high accuracy and productively contributes to the healthcare system.

REFERENCES

[1] A. Myers, "Thyroid Disorders Item Type Thesis." [Online]. Available: http://hdl.handle.net/10484/12100

[2] M. Mohamedali, S. Reddy Maddika, A. Vyas, V. Iyer and P. Cheriyath, "Thyroid disorders and chronic kidney disease," *International Journal of Nephrology*, vol. 2014, 2014. doi: 10.1155/2014/520281.

[3] N. Shah, "Prevalence and Manifestations of Hypothyroidism among Population of Hyderabad, Sindh, Pakistan," *Pure and Applied Biology*, vol. 10, no. 2021, doi: 10.19045/bspab.2021.100069.

[4] K. Salman and E. Sonuc, "Thyroid Disease Classification Using Machine Learning Algorithms," *Journal of Physics*, 2021. doi: 10.1088/1742-6596/1963/1/012140.

[5] T. Alyas, K. Alissa, A. S. Mohammad, S. Asif, T. Faiz et al., "Innovative fungal disease diagnosis system using convolutional neural network," *Computers, Materials & Continua*, vol. 73, no.3, pp. 4869–4883, 2022.

[6] N. Tabassum, A. Rehman, M. Hamid, M. Saleem, S. Malik et al., "Intelligent nutrition diet recommender system for

diabetic's patients," *Intelligent Automation & Soft Computing*, vol. 30, no.1, pp. 319–335, 2021.

[7] S. Agrebi and A. Larbi, "Use of artificial intelligence in infectious diseases," in Artificial Intelligence in Precision Health, *Elsevier*, pp. 415–438, 2020. doi: 10.1016/b978-0-12-817133-2.00018-5.

[8] K. Guleria, S.Sharma, S. Kumar and S.Tiwari, Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning, Measurement: *Sensors*, Vol. 24,2022.

[9] G. Chaubey, D. Bisen, S. Arjaria and V. Yadav, "Thyroid Disease Prediction Using Machine Learning Approaches," *National Academy Science Letters*, vol. 44, no. 3, pp. 233–238, 2021, doi: 10.1007/s40009-020-00979-z.

[10] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz and H. J. W. L. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, Aug. 01, 2018. doi: 10.1038/s41568-018-0016-5.

[11] E. J. Ha and J. H. Baek, "Applications of machine learning and deep learning to thyroid imaging: Where do we stand?," *Ultrasonography*, vol. 40, no. 1, pp. 23–29, 2021. doi: 10.14366/usg.20068.

[12] B. Zhang, "Machine Learning-Assisted System for Thyroid Nodule Diagnosis," *Thyroid*, vol. 29, no. 6, pp. 858–867, 2019, doi: 10.1089/thy.2018.0380.

[13] T. Alyas, M. Hamid, K. Alissa,T. Faiz, N. Tabassum and A. Ahmad, Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach, *BioMed Research International*, vol. 2022, no. 22, pp. 1-15, 2022. doi: 10.1155/2022/9809932.

[14] S. Borzouei, H. Mahjub, N. Sajadi and M. Farhadian, "Diagnosing thyroid disorders: Comparison of logistic regression and neural network models," *Journal of Family Medicine and Primary Care*, vol. 9, no. 3, p. 1470, 2020, doi: 10.4103/jfmpc.jfmpc_910_19.

[15] B. Li and LP. Zhao, Multiconditional machining process quality prediction using deep transfer learning network. *Advances in Manufacturing*, vol. 11, pp. 329–341, 2023. https://doi.org/10.1007/s40436-022-00415-z

[16] F. Temurtas, "A comparative study on thyroid disease diagnosis using neural networks," *Expert System Application*, vol. 36, no. 1, pp. 944–949, 2009, doi: 10.1016/j.eswa.2007.10.010.

[17] M. A. A. R. Asif , "Computer aided diagnosis of thyroid disease using machine learning algorithms," in *Proceedings of 2020 11th International Conference on Electrical and Computer Engineering, ICECE 2020, Institute of Electrical and Elec-tronics Engineers Inc.*, Dec. 2020, pp. 222–225. doi: 10.1109/ICECE51571.2020.9393054.

[18] H. Rehman, C. Y. Lin, Z. Mushtaq and S. F. Su, "Performance Analysis of Machine Learning Algorithms for Thyroid Disease," *Arabian Journal for Science and Engineering*, vol. 46, no. 10, pp. 9437–9449, 2021, doi: 10.1007/s13369-020-05206-x.

[19] Duggal, Priya and Shipra Shukla. "Prediction Of Thyroid Disorders Using Advanced Machine Learning Techniques." *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* pp. 670-675, 2020.

[20] N. Uzir, S. Raman, S. Banerjee, and R. S. Nishant Uzir Sunil R, "Experimenting XGBoost Algorithm for Prediction and Clas-sification of Different Datasets Experimenting XGBoost Algorithm for Prediction and Classifi cation of Different Datasets," *International Journal of Control Theory and Applications*, vol. 9, 2016.

[21] G. Abdurrahman and M. Sintawati, "Implementation of xgboost for classification of parkinson's disease," in Journal of Physics, vol 23, 2020. doi: 10.1088/1742-6596/1538/1/012024.

[22] G. Battineni, N. Chintalapudi and F. Amenta, "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)," *Informatics in Medicine Unlocked*, vol. 16, 2019, doi: 10.1016/j.imu.2019.100200.

[23] P. Cunningham and S. J. Delany, "K-Nearest Neighbour Classifiers-A Tutorial," *ACM Computing Surveys*, vol. 54, no. 6, 2021. doi: 10.1145/3459665.

[24] B. Mahesh, "Machine Learning Algorithms-A Review Machine Learning Algorithms-A Review View project Self Flowing Generator View project Batta Mahesh Independent Researcher Machine Learning Algorithms-A Review," *International Journal of Science and Research*, 2018, doi: 10.21275/ART20203995.

[25] T. Rymarczyk, E. Kozłowski, G. Kłosowski and K. Niderla, "Logistic regression for machine learning in process tomography," *Sensors*, vol. 19, no. 15, Aug. 2019, doi: 10.3390/s19153400.

[26] V. Nasteski, "An overview of the supervised machine learning methods," *horizons*, vol. 4, pp. 51–62, 2017, doi: 10.20544/horizons.b.04.1.17.p05.

[27] E. Sevgen, S. Kocaman, H. A. Nefeslioglu and C. Gokceoglu, "A novel performance assessment approach using photo-grammetric techniques for landslide susceptibility mapping with logistic regression, ann and random forest," *Sensors*, vol. 19, no. 18, Sep. 2019, doi: 10.3390/s19183940.

IEER
Institute for Excellence in Education & Research