



ORIGINAL

## Early detection of thyroid disease using feature selection and hybrid machine learning approach

### Detección temprana de enfermedades tiroideas mediante selección de características y un enfoque híbrido de aprendizaje automático

Barnokhon Badridinova<sup>1</sup>  , Kamola Azimova<sup>2</sup>  , Gulnoza Iskandarova<sup>3</sup>  , Gulruh Majidova<sup>4</sup>  , Xasan Abdullaev<sup>5</sup>  , Muso Urinov<sup>6</sup>  , Farida Tokhirova<sup>7</sup>  

<sup>1</sup>Bukhara State Medical Institute Named After Abu Ali Ibn Sino, Uzbekistan.

<sup>2</sup>Samarkand State Medical University, Uzbekistan.

<sup>3</sup>Center for the Development of Professional Qualifications of Medical Workers of the Ministry of Health of the Republic of Uzbekistan, Uzbekistan.

<sup>4</sup>Jizzakh State Pedagogical University, Uzbekistan.

<sup>5</sup>Samarkand State University, Samarkand, Uzbekistan.

<sup>6</sup>Bukhara State Medical Institute, Uzbekistan.

<sup>7</sup>Samarkand State Medical University, Uzbekistan.


**Cite as:** Badridinova B, Azimova K, Iskandarova G, Majidova G, Abdullaev X, Urinov M, et al. Early detection of thyroid disease using feature selection and hybrid machine learning approach. Health Leadership and Quality of Life. 2024;3:192. <https://doi.org/10.56294/hl2024.192>

Submitted: 01-03-2024

Revised: 20-06-2024

Accepted: 15-10-2024

Published: 16-10-2024

Editor: PhD. Prof. Neela Satheesh 

Corresponding author: Barnokhon Badridinova 

#### ABSTRACT

In today's environment, thyroid disorders are quite widespread and widely dispersed. They frequently result in serious physical and mental suffering. It interferes with the thyroid gland's ability to operate, which causes the thyroid to secrete too much hormone. The thyroid organs are ground up by the hormones produced when the body enters auto-safe mode in this illness. Avoiding this condition is crucial because it has irreversible effects on the body. Since this disorder is extremely difficult to cure once it reaches its final stage, preventing it from occurring needs some awareness of its development. The ontological challenges and disparate data standards that are employed in Medical Data Analysis (MDA) and system-assisted healthcare management are well-known in the healthcare industry. Rapid technological breakthroughs have drawn researchers to the health sector to create accurate, dependable, and reasonably priced medical (DSS) decision support systems (MDSS). Therefore, there is continuous research being done to construct an efficient and practically applicable MFFN+MLP-based DSS for medical data (MD) processing and knowledge discovery (KD). Using computerised intelligent medical decision support systems offers a practical way to help medical professionals diagnose patients quickly and correctly. Before a practical medical diagnosis system can be created and implemented, a number of problems must be addressed and handled, including how to make decisions when faced with ambiguity and imprecision.

**Keywords:** Cervical Disease; Computer-Aided Diagnostic (CAD) Systems; Machine Learning (ML); Decision Tree (DT).

#### RESUMEN

En el entorno actual, los trastornos de la tiroides están muy extendidos y ampliamente dispersos. Con frecuencia provocan un sufrimiento físico y mental grave. Interfieren en la capacidad de funcionamiento de la glándula tiroides, lo que hace que la tiroides segregue demasiada hormona. Los órganos tiroideos se

desgastan por las hormonas producidas cuando el cuerpo entra en modo de seguridad automática en esta enfermedad. Evitar esta afección es crucial porque tiene efectos irreversibles en el cuerpo. Dado que este trastorno es extremadamente difícil de curar una vez que llega a su etapa final, para evitar que se produzca es necesario tener cierta conciencia de su desarrollo. Los desafíos ontológicos y los estándares de datos dispares que se emplean en el análisis de datos médicos (MDA) y la gestión de la atención médica asistida por sistemas son bien conocidos en la industria de la salud. Los rápidos avances tecnológicos han atraído a los investigadores al sector de la salud para crear sistemas de apoyo a la toma de decisiones médicas (MDSS) precisos, confiables y a un precio razonable. Por lo tanto, se están realizando investigaciones continuas para construir un DSS basado en MFFN + MLP eficiente y prácticamente aplicable para el procesamiento de datos médicos (MD) y el descubrimiento de conocimiento (KD). El uso de sistemas informáticos inteligentes de apoyo a la toma de decisiones médicas ofrece una forma práctica de ayudar a los profesionales médicos a diagnosticar a los pacientes de forma rápida y correcta. Antes de poder crear e implementar un sistema práctico de diagnóstico médico, se deben abordar y manejar una serie de problemas, incluido cómo tomar decisiones cuando se enfrentan a la ambigüedad y la imprecisión.

**Palabras clave:** Enfermedad Cervical; Sistemas de Diagnóstico Asistido por Ordenador (CAD); Aprendizaje Automático (ML); Árbol de Decisiones (DT).

## INTRODUCTION

As per the survey, around one in every 38 000 individuals worldwide are affected by congenital hypothyroidism. Thyroid disorders affect almost 42 million people in poor nations like India. According to data, thyroid diseases are on the rise in India. About ten percent of Indian adults have hypothyroidism. The inability of the thyroid gland to generate enough thyroid hormones to fulfil the body's needs is known as hypothyroidism. This syndrome is widespread among women of childbearing age and is twice as common in women as in men.<sup>(1)</sup> According to a 2016 study carried out in nine Indian states, 13.13 percent of pregnant women had hypothyroidism based on TSH levels. Indians appear to be more likely to experience it; in Mumbai, the incidence is one out of 2640. These days, more than 25 000 hospitals worldwide gather patient data in a variety of formats. Clinical and medical research are conducted utilising statistical tests and classical analysis in the traditional way.<sup>(2)</sup> Human hormone fluctuation is the cause of the thyroid, and it is crucial for humans to maintain a balanced hormonal level despite this minimal fluctuation. Hormonal disruption includes some risk aspects therefore it is more essential to continue concern. and determine the appropriate diagnosis at the appropriate time for the physicians. These days, expert advisory systems (EAS) can be provided by computer science specialists thanks to advancements in information and technology in the medical sciences.<sup>(3)</sup> In this situation, maintaining good health has become extremely difficult for people. However, as the medical field develops quickly, new technologies are created that make it easier for doctors to use machine learning techniques to precisely identify disorders. The conditions that affect the thyroid gland, which is a butterfly-shaped gland in front of the neck, are known as thyroid diseases. It plays a crucial function in controlling the various metabolic processes that occur throughout the body. The term "isthmus" refers to the thin region of tissue that sits between the glands. It connects the two thyroid lobes on either side.<sup>(4)</sup> Thyroid gland function is controlled by a brain-based feedback mechanism. The brain's hypothalamus creates a hormone called thyrotrophic, which can be used to release thyroid hormone when levels are low. Thyroid stimulating hormone is released by the pituitary gland, which is situated at the base of the brain. The pituitary gland and hypothalamic diseases regulate the thyroid gland. Thyroid disorders can also be brought on by these tissues' influence on thyroid function.<sup>(5)</sup> Machine learning (ML) techniques have been applied to improve thyroid prediction outcomes by different sectors and businesses, following the lead of successful applications in medical diagnostics, bioinformatics, brain-machine interfaces, e-business, marketing, and retail. ML is being utilised in the healthcare sector to identify fraudulent claims in the medical insurance market. ML can benefit the health care industry in a number of ways, despite the fact that it carries some risk. The list of applications for ML in the healthcare sector is provided below.<sup>(6)</sup>

The next section of paper is arranged as follows: Section 2 observing about many researchers to develop the models to diagnose the thyroid disorder. Section 3 proposed a novel methodology for creating an aid system thyroid diagnosis. Section 4 describes the experiment analysis. And section 5 presents the work's conclusion. Anatomy of thyroid and parathyroid gland is given in figure 1.

## Related work

Numerous scholars have employed diverse forms of Data Mining (DM) techniques.<sup>(7,8,9,10,11,12)</sup> Future work involving several datasets and algorithms linked to this work will demonstrate that an adequate approach and assurance are obtained in identifying thyroid problems. The prediction of thyroid diagnosis has emerged as a reputable field in ML research, with AI-based methods employed to address the intricate problem of thyroid illness classification. (table 1).

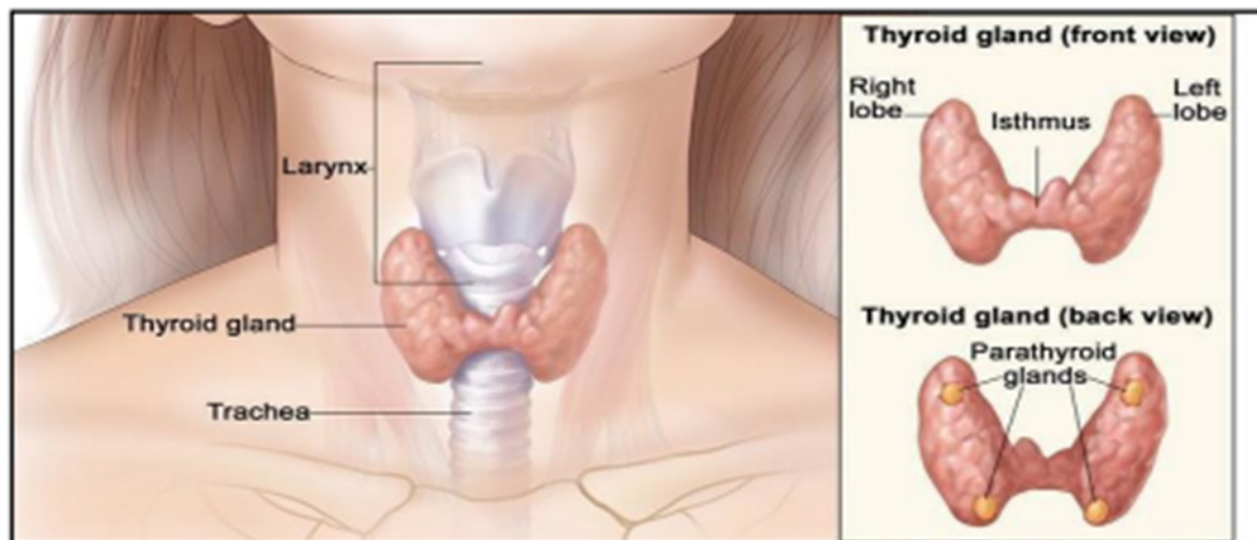


Figure 1. Anatomy of Thyroid and parathyroid Gland

Table 1. Some of the related works

S.No	Paper And Author Details	Objective	Method
1	An Effective ML Method for Classifying the Glyphosate Poisoning Status in Rats Using Blood Routine Test", Jiayin Zhu ; Xuehua Zhao ; Huaizhong Li ; Huiling Chen, 2018	It suggests a fresh method for modelling glyphosate poisoning using blood samples, an innovative and efficient technique for automatically determining the glyphosate poisoning status, and the ability to pick the most pertinent indices with the help of a meta-heuristic (MH)-based feature.	PSO-FKNN, FKNN
2	"Identification of thyroid nodules in infrared images by CNN", M. B. H. Moran ; A. Conci ; J. R. González ; A. S. Araújo ; W. G. Fiirst ; Charbel P. Damião ; Giovanna A. B, 2018	In this, we described a system that uses CNN (convolutional neural networks) and basic image processing techniques to locate thyroid nodules in thermograms.	CNN.
3	High-Resolution Raman Microscopic Detection of Follicular Thyroid Cancer Cells with Unsupervised ML, J. Nicholas Taylor et al., 2019.	Discover that, in comparison to Nthy-ori 3-1, malignant FTC133 cells have higher populations of lipid-carrying components and lower populations of cytochrome-containing components. Additionally, the regions harbouring these contributions are primarily located outside of the cell nucleus.	This method not only improves classification accuracy but also offers extensive subcellular information regarding the biochemical distinctions and cellular locations between normal and malignant follicular thyroid cells.
4	Thyroid cancer diagnosis by Raman spectroscopic/ Macroscopic, et al., 2020.	Reliable cancer diagnosis can be obtained by identifying thyroid tissues using the Raman Spectroscope. It making use of several algorithms with the confusion matrix (CM). 45 medical datasets serve as samples.	Enrolment, clinical assessment, neck ultrasonography, typical tissue, preparation, and Raman spectroscopy collection cluster analysis were all studied. 91 % accuracy in biomedical analysis is considered
5	Malignant and benign thyroid nodule differentiation through the analysis of blood plasma with terahertz spectroscopy/MARIA R. KONNIKOVA, et al., 2021.	There was a correlation between the samples and absorption at 1 THz and the content of glucose and the level of miRNA-146b. For the analysis of THz spectra, an ensemble approach with two stages was suggested.	Use a linear kernel (SVM) support vector machine to distinguish between participants with thyroid nodules and those without. In the second step, participants with benign and malignant thyroid nodules were separated using a machine using additional data preprocessing using the Ornstein-Hollenbeck kernel Principal Component Analysis (PCA). Learning displayed

## Objectives

The ability to handle noisy and missing data, transparency of diagnostic knowledge with the ability to explain the choice, and the ability to minimise the number of tests needed for an excellent diagnosis, are some of the desired features from the ML system for medical diagnosis.

To find out if a patient is accepting the existence of the illness, an integrated framework-based machine learning technique was created. The study is focused on thyroid disease due to the inherent ambiguity associated with diagnosing the condition and treating it.

Although the thyroid gland may operate normally despite initial morphological alterations, a thorough examination is necessary to assess the structural and functional aspects of the gland. It is very difficult for doctors to determine whether thyroid disease would manifest at a structural level or not during the diagnostic phase because of the misleading nature of normal readings with slight deviations. Making the appropriate gland assessment would therefore prevent a patient from receiving a false diagnosis, which is one of the main issues facing the healthcare sector.

## Proposed Framework

The primary goal of this study was to develop an integrated model for thyroid illness prediction utilising a competitive ML algorithm. In particular, this research develops an integrated framework model and validates it with an existing model, all while using recognised variables (input and output) for thyroid illness diagnosis (based on published research). The most popular diagnostic assessment in endocrine medicine is thyroid function testing, which is used as a screening tool, to confirm the clinical diagnosis of hyper- and hypothyroidism, to gauge how well a patient is responding to treatment, and to monitor patients with differentiated thyroid cancer. By using an integrated model, it may be possible to forecast how patients and physicians would handle thyroid disease with care and provide comprehensive recommendations for societal development. categorising potential factors (input and output) that may have an impact on the diagnosis of thyroid illness and looking at the connections between them.<sup>(13)</sup>

Preprocessing is done on the data obtained from the UCI repository, where the masking method is used to check for missing values rather than numerical constraints. The column mean value is used in place of the missing value or Not a Number (NaN) values if they exist. Differential Evolution (DE), a hybrid algorithm, receives the preprocessed data. A subset of the child records is created from the parent records using this procedure. To assess fitness, the subsets of data are fed into the Kernel Based Bayesian classifier algorithm. Error stabilisation is used to gauge fitness. Following stabilisation, the results are divided into three categories: normal, hyperthyroid, and hypothyroid.

## Data acquisition

The information is gathered from multiple sources. Next, the gathered data is normalised. The dataset is then divided into training and testing subsets. There are two possible formats for the patient data: structured and unstructured. To analyse the data, the unstructured data must be converted into structured data. The dataset's input, which includes three categories Normal, hypothyroid, hyperthyroid, and normal. Figure 2 represents the framework of existing framework.

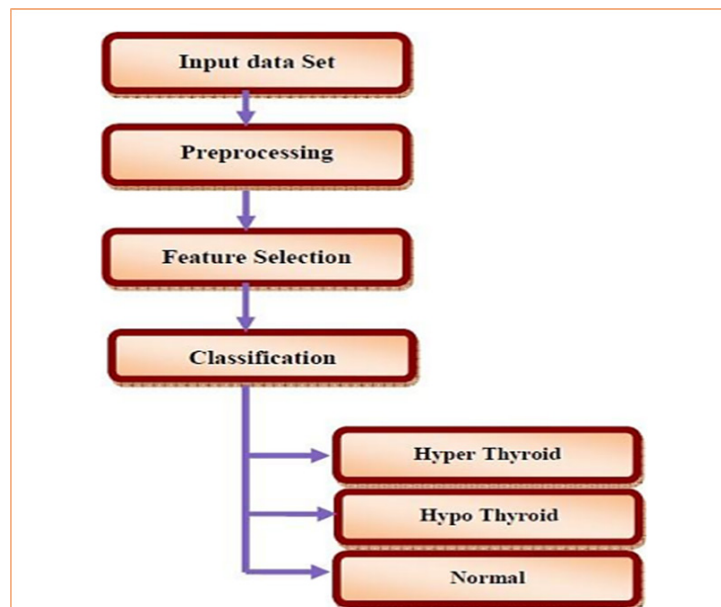


Figure 2. Framework of Existing Framework

## Preprocessing

To address a variety of issues, such as noisy data, redundant data, missing data values, etc., pre-processing is required. High-quality data will produce high-quality outcomes and lower machine learning expenses. It has an impact on the outcomes. Preprocessing is required before using any DM techniques in order to increase the effectiveness and mining process for the thyroid illness dataset. The mean value will be used to fill in the blanks. Preprocessing missing data is necessary to enable the necessary algorithm to process the entire data set. Furthermore, data sets with discrete features may collect knowledge from most existing methods. Preprocessing is given in figure 4.

The methods can be merged to provide discrete attributes if the features are continuous. Both continuous and discrete data are preprocessed in the proposed study using data obtained from blood samples gathered from thyroid patient's repository. Using the masking technique, the missing value and not a numerical constraint are checked at this point. The column mean value is used in place of the missing value or Not a Number (NaN) values if they exist. Preprocessing is a crucial step in data mining that transforms incomplete data into a legible format. Data load from the repository is illustrated in figure 3.

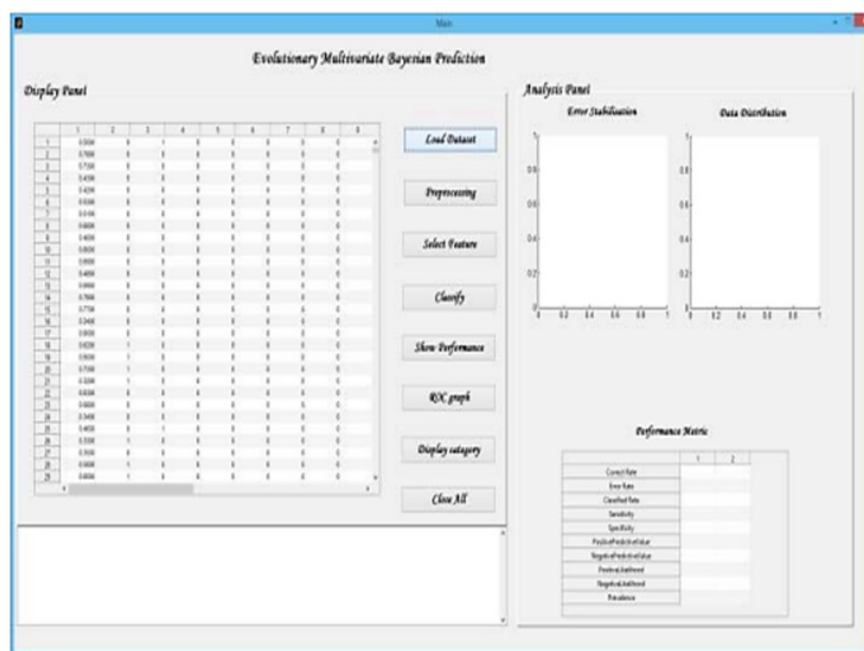


Figure 3. Data Load from the repository

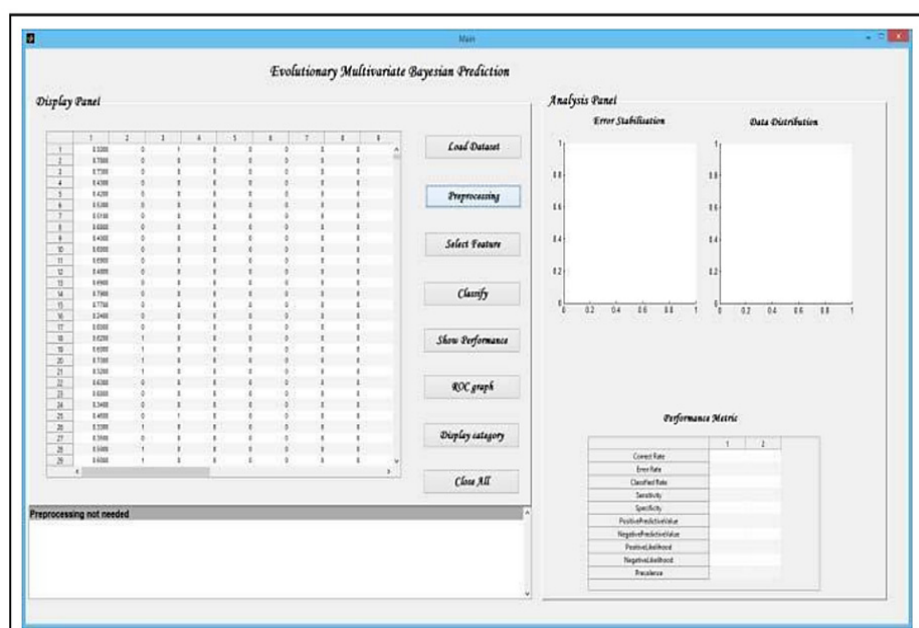


Figure 4. Preprocessing



Feature selection (FS) is another name for attribute selection in ML. It is a method for selecting pertinent elements and streamlining models in order to better comprehend consumers. It is intended to cut down on training time. It's not the same as feature extraction (FE). Using the correlation attribute evaluation method, the dimension of the thyroid dataset was reduced. The relation between each property is computed using it. After identifying the pertinent attributes with a moderate to positive or negative association, eliminate the attribute values that are closer to zero. Lastly, to accept the chosen qualities. Precise disease diagnosis and effective treatment are essential components of the priceless medical care that the health-care system provides for its patients. Medical databases are distinct due to their varied, massive, and privacy-sensitive data, which present data mining issues.<sup>(14)</sup>

### Hybrid Multilayer Feed Forward Neural Network (MFFNN) with Multi-Layer Perceptron (MLP)

Designing a better thyroid disease pattern classifier system with other parameters including age group, heredity, and antibodies is necessary to fill in the significant research gaps. The best way to handle this issue is to take into account other factors, as the blood test is the worst and most basic way to determine whether or not a person has thyroid issues. Furthermore, the system might make use of stronger classifiers to raise the diagnostic system's overall accuracy. The most recent ML methods must be employed by the upgraded thyroid system for training and evaluating the ML model. The comprehensive models are provided in this part.<sup>(15)</sup>

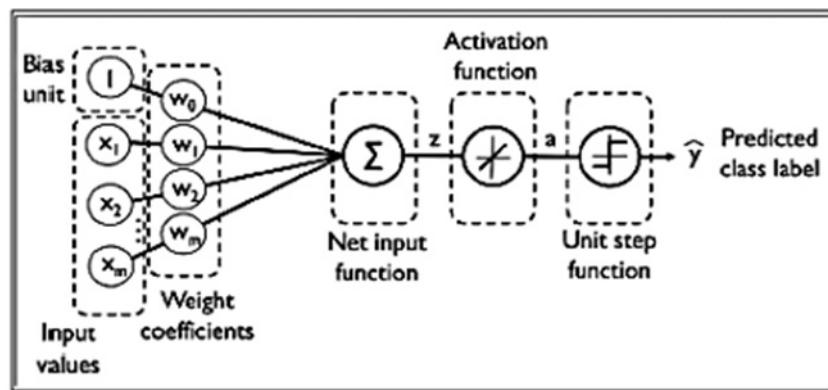


Figure 5. MFFNN

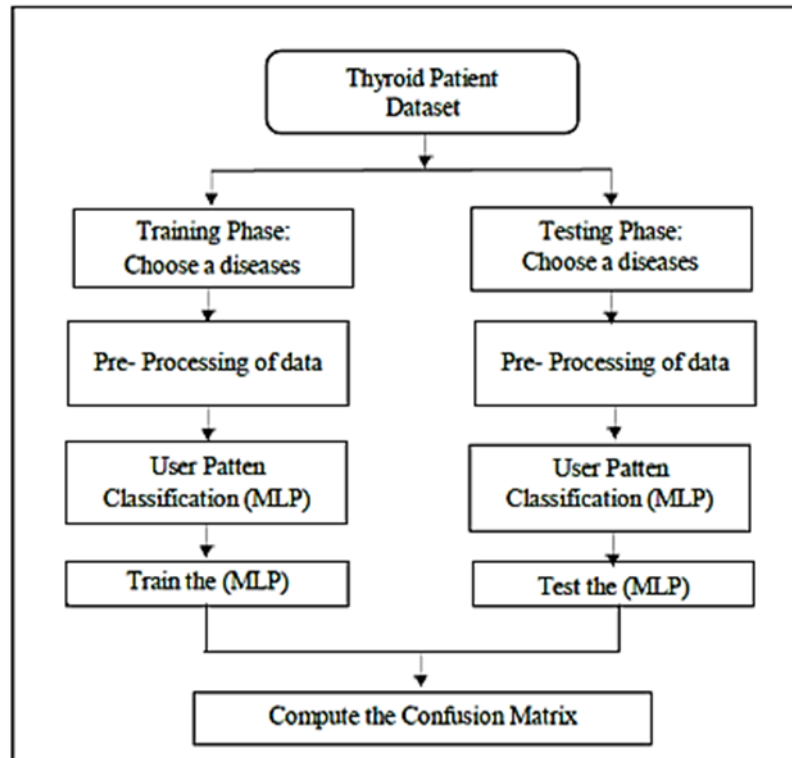


Figure 6. Frame work of Training and Test of MLP Model

A MFFNN is composed of an Output Layer (OL), one or more Hidden Layers (HL), and an Input Layer (IL).<sup>(16)</sup> Figure 5 illustrates a multilayer feed-forward network. Units make up each stratum. The network's inputs line up with the characteristics that are measured for every training pair.<sup>(17)</sup> The units that comprise the IL receive the inputs concurrently.<sup>(18)</sup> After passing through the IL, these inputs are simultaneously supplied to a HL, which is a second layer of "neuronlike" units, after being weighted.<sup>(19)</sup> One HL outputs can be fed into another HL outputs, and so on. Although in reality, only one HL is typically employed, the number of HL is variable.<sup>(20)</sup> The network's prediction for the given tuples is produced by the OL units, which receive as input the weighted outputs of the final HL.<sup>(21)</sup>

The units in the IL are known as input units. The units in the HL and OL are also commonly called neurodes or output units due to their biological basis.<sup>(22)</sup> The multilayer (NN) neural network seen in figure 5 has two layers of output units. Because of this, we call it a two-layer NN.<sup>(23)</sup> (The IL is ignored since its only function is to transfer the input values to the subsequent layer.) Similarly, a NN with two HL is called a three-layer NN, and so on.<sup>(24)</sup> The network is feed-forward since none of the weights cycle back to an input unit or an output unit from a previous layer.<sup>(25)</sup> The following forward layer is Fully Connected (FC) since each unit receives input from every other unit. Figure 6 represents the frame work of training and test of MLP model.<sup>(26)</sup>

To assign a particular patient to one of the three classes, a dataset of 120 samples has been obtained and preprocessed.<sup>(27)</sup> The data collection is manually enriched to exclude noise, abnormalities, and Boolean values. After the dataset is ready, it is used to develop and train a multilayer pattern classification model. The MLP pattern classifier model is retained for the testing phase.<sup>(28)</sup> To evaluate the accuracy of thyroid predictions, a sample of randomly chosen patients is subjected to the stored MLP prediction system.

## RESULTS AND DISCUSSION

The trials were conducted using a real dataset of 120 cases that were collected from SKIMS Soura in Srinagar, India.<sup>(29)</sup> The subjects were carefully selected to represent a broad spectrum of the population, including males, women, the elderly, and children. For each of the 120 examples, the values for eleven different qualities were gathered, and some of the attributes have even been factored and quantified.<sup>(30)</sup> The dataset has undergone preprocessing to get rid of mistakes, abnormalities, and ambiguities.<sup>(31)</sup> The pre-processed dataset is used to train the pattern classification model utilising the back error propagation strategy for the MLP.<sup>(32)</sup>

Train and test data instances have been separated from the original data set. This was accomplished by altering the training dataset's size and then testing the ML model on a test data set.<sup>(33)</sup> Comparison of different phases of processing is given in table 2.

Iteration No	Data Size	Performance	Accuracy			
			Training	Validation	Testing	Overall
1	30	0,031	85	20	40	66,70
2	60	9,87	98,25	77	98,50	99,10
3	90	0,000158	99,12	99	99,01	99,50
4	120	3,1	99,23	99	99,20	99,7

The objective of this work was to determine which common ML techniques are more accurate for the Human Thyroid Disorder Prediction using data gathered from the UCI repository. The first research question's response is to identify the method that performs the best.<sup>(34)</sup> The second research question focusses more on data preparation methods. Although these two study questions appear to be focused on unrelated topics, they are actually closely related when considering the effectiveness of the techniques in table 3.

Algorithm	Accuracy	Precision	Recall	F1 score
ANN	88,4	88,1	92	0,884
KNN	88	87	91,7	0,88
DT	87	80	88,4	0,897
LR	88	82,2	90	0,88
MFFN+MLP	98,7	98	98	0,97

The model's accuracy, precision, recall, and F1score were also used to assess its performance. The results

demonstrate that the Prediction classifier model, which is a modified decision tree method, gets a decent outcome. Ten assessment metrics are used to assess the outcomes, and accuracy-which has an accuracy of 98,7% and precision of 98-provides the most accurate and precise results.

## CONCLUSIONS

Thyroid hormone problems are a serious issue for many. With its autonomous forecasts and prompt patient treatment, the thyroid prediction employing machine learning techniques opened up new avenues for technological advancement. The training data is used to construct the classification model, and during the prediction step, the data is put to the test through predictions. The thyroid datasets are tested by the implementation of several machine algorithms. When compared to other methods, the hybrid MFFN+MLP algorithm produced the best accuracy results when it came to solving classification and regression problems. This makes it abundantly evident that healthcare organizations can use classifying algorithms to anticipate technologies for healthy living.

## REFERENCES

1. Olatunji, Sunday O., Sarah Alotaibi, Ebtisam Almutairi, Zainab Alrabae, Yasmeen Almajid, Rahaf Altabee, Mona Altassan, Mohammed Imran Basheer Ahmed, Mehwash Farooqui, and Jamal Alhiyafi. "Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset." *Computers in biology and medicine* 131 (2021): 104267.
2. Chang, Chuan-Yu, Shao-Jer Chen, and Ming-Fong Tsai. "Application of support-vector-machine-based method for feature selection and classification of thyroid nodules in ultrasound images." *Pattern recognition* 43, no. 10 (2010): 3494-3506.
3. Wu, Dan, Yan Zhuang, Guoliang Liao, Lin Han, Ke Chen, Cheng Li, Zhan Hua, and Jiangli Lin. "Feature Selection and Classification Technique for Predicting Lymph Node Metastasis of Papillary Thyroid Carcinoma." *Journal of Computational Biophysics and Chemistry* 23, no. 06 (2024): 801-814.
4. Misra, Puneet, and Arun Singh Yadav. "Improving the classification accuracy using recursive feature elimination with cross-validation." *Int. J. Emerg. Technol* 11, no. 3 (2020): 659-665.
5. Bhende, Deepali, Gopal Sakarkar, Suhasini Chaurasia, Neetu Amlanc, Swapnil Deshpande, Priyanka Samarth, and Zohra Yasmeen. "Machine Learning-Based Classification of Thyroid Disease: A Comprehensive Study on Early Detection and Risk Factor Analysis." In *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1-6. IEEE, 2024.
6. Chaganti, R., F. Rustam, I. De La Torre Díez, J. L. V. Mazón, C. L. Rodríguez, and I. Ashraf. "Thyroid Disease Prediction Using Selective Features and Machine Learning Techniques. *Cancers* 2022, 14, 3914." (2022).
7. Obaido, George, Okechinyere Achilonu, Blessing Ogbuokiri, Chimeremma Sandra Amadi, Lawal Habeebullahi, Tony Ohalloran, Chidozie Williams Chukwu et al. "An Improved Framework for Detecting Thyroid Disease Using Filter-Based Feature Selection and Stacking Ensemble." *IEEE Access* (2024).
8. Latif, Muhammad Armghan, Zohaib Mushtaq, Saad Arif, Sara Rehman, Muhammad Farrukh Qureshi, Nagwan Abdel Samee, Maali Alabduhafith, Yeong Hyeon Gu, and Mohammed A. Al-masni. "Improving Thyroid Disorder Diagnosis via Ensemble Stacking and Bidirectional Feature Selection." *Computers, Materials & Continua* 78, no. 3 (2024).
9. Akhtar, Tehseen, Saad Arif, Zohaib Mushtaq, Syed Orner Gilani, Mohsin Jamil, Yasar Ayaz, and Shahid Ikramullah Butt. "Ensemble-based effective diagnosis of thyroid disorder with various feature selection techniques." In *2022 2nd International Conference of Smart Systems and Emerging Technologies (SMARTTECH)*, pp. 14-19. IEEE, 2022.
10. Shiuh, Tong Lim, Wah Khaw Khai, Ying Chew Xin, and Chung Yeong Wai. "Prediction of Thyroid Disease using Machine Learning Approaches and Featurewiz Selection." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 15, no. 3 (2023): 9-16.
11. Sultana, Azrin, and Rakibul Islam. "Machine learning framework with feature selection approaches for thyroid disease classification and associated risk factors identification." *Journal of Electrical Systems and Information Technology* 10, no. 1 (2023): 32.



12. Tutsoy, Onder, and Hilmi Erdem Sumbul. "A novel deep machine learning algorithm with dimensionality and size reduction approaches for feature elimination: thyroid cancer diagnoses with randomly missing data." *Briefings in Bioinformatics* 25, no. 4 (2024).
13. Riajuliislam, Md, Khandakar Zahidur Rahim, and Antara Mahmud. "Prediction of thyroid disease (hypothyroid) in early stage using feature selection and classification techniques." In *2021 International conference on information and communication technology for sustainable development (ICICT4SD)*, pp. 60-64. IEEE, 2021.
14. Duggal, Priyanka, and Shipra Shukla. "Prediction of thyroid disorders using advanced machine learning techniques." In *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 670-675. IEEE, 2020.
15. Chaganti, Rajasekhar, Furqan Rustam, Isabel De La Torre Díez, Juan Luis Vidal Mazón, Carmen Lili Rodríguez, and Imran Ashraf. "Thyroid disease prediction using selective features and machine learning techniques." *Cancers* 14, no. 16 (2022): 3914..
16. Maja, O., Mirko, B., Goran, M., & Vladimir, P. M. (2019). The Influence of Ocean Tides to Determine the Earth's Orientation Parameters. *Archives for Technical Sciences*, 2(21), 43-53.
17. Ergüden, D., Gürlek, M., Kabaklı, F., & Turan, C. (2022). The First Occurrence of Xanthochromic Fish, *Diplodus sargus* (Family: Sparidae) in the Eastern Mediterranean. *Natural and Engineering Sciences*, 7(1), 34-40. <http://doi.org/10.28978/nesciences.1098658>
18. Kamarudin, M. S., Nuruljannah, M. P., Syukri, F., & Cruz, C. R. (2023). Effects of dietary protein-energy level on the survival, growth and body composition of tinfoil barb, *Barbonymus schwanenfeldii* fry. *International Journal of Aquatic Research and Environmental Studies*, 3(2), 35-50. <https://doi.org/10.70102/IJARES/V3I2/3>
19. Buljubašić, S. (2020). Application of New Technologies in the Water Supply System. *Archives for Technical Sciences*, 1(22), 27-34.
20. Sumithra, S., & Sakshi, S. (2024). Exploring the Factors Influencing Usage Behavior of the Digital Library Remote Access (DLRA) Facility in a Private Higher Education Institution in India. *Indian Journal of Information Sources and Services*, 14(1), 78-84. <https://doi.org/10.51983/ijiss-2024.14.1.4033>
21. Dülger, G., & Dülger, B. (2022). Antibacterial Activity of *Stachys sylvatica* Against Some Human Eye Pathogens. *Natural and Engineering Sciences*, 7(2), 131-135. <http://doi.org/10.28978/nesciences.1159224>
22. Prabadevi, M. N., Mary Auxilia, P. A., Subramanian, K. P, & Rengarajan, V. (2024). Strategies for Leveraging Digital Libraries to Improve Financial Literacy among Rural Entrepreneurial Women. *Indian Journal of Information Sources and Services*, 14(2), 28-33. <https://doi.org/10.51983/ijiss-2024.14.2.05>
23. Makhloogh, A., Nasrollahzadeh Saravi, H., Naderi, M. J., Eslami, F., & Ahmadnezhad, A. (2023). Use of algal indices for determining of water quality in the Sirvan River tributaries (Kurdistan-Iran). *International Journal of Aquatic Research and Environmental Studies*, 3(1), 43-56. <https://doi.org/10.70102/IJARES/V3I1/5>
24. Ćurčić, M., Milinković, Dragica, Radivojević, D., & Đurić, Dijana. (2019). Vertical Distribution of Diatoms on Mosses in Wells of Bijeljina Municipality in Bosnia and Herzegovina. *Archives for Technical Sciences*, 1(20), 53-64.
25. Demirci, B., & Demirhan, S. A. (2022). Food composition and dietary overlap of the lionfish species in Iskenderun Bay. *Natural and Engineering Sciences*, 7(3), 228-239. <http://doi.org/10.28978/nesciences.1163001>
26. Sathish Kumar, M., Santhi, L., & Senthilkumar, A. (2024). Quantifying the Impact of Indian Virtual Reality Research: A Scientometric Study. *Indian Journal of Information Sources and Services*, 14(3), 1-5. <https://doi.org/10.51983/ijiss-2024.14.3.01>
27. Armnazi, M, & Alegasan, M. (2024). Target Situation Needs Analysis of English Language Skills Required by Syrians in the Arabian Gulf Area. *Indian Journal of Information Sources and Services*, 14(3), 77-85. <https://doi.org/10.51983/ijiss-2024.14.3.11>

28. Admodisastro, V. A., Ransangan, J., Ilias, N., & Tan, S. H. (2022). Oyster farming potential in Sabah, Malaysia. *International Journal of Aquatic Research and Environmental Studies*, 2(1), 17-22. <https://doi.org/10.70102/IJARES/V2I1/3>
29. Radmanović, S., Đorđević, A., & Nikolić, N. (2018). Humus Composition of Rendzina Soils in Different Environmental Conditions of Serbia. *Archives for Technical Sciences*, 2(19), 57-64.
30. Saç, G., & Özüluğ, M. (2021). Length-weight and length-length relationships of three endemic freshwater fish species from Anatolia, Turkey. *Natural and Engineering Sciences*, 6(1), 53-59. <http://doi.org/10.28978/nesciences.868080>
31. Salokhiddinov, A. T., et al. (2020). Climate change effects on irrigated agriculture: Perspectives from agricultural producers in eastern Uzbekistan. *IOP Conference Series: Earth and Environmental Science*, 612, 012058. <https://doi.org/10.1088/1755-1315/612/1/012058>
32. Karimov, B. K., et al. (2020). Relationship between the concentrations of nitrogen compounds and the water discharge in the Chirchik River, Uzbekistan. *IOP Conference Series: Earth and Environmental Science*, 614, 012154. <https://doi.org/10.1088/1755-1315/614/1/012154>
33. Karimov, N., et al. (2024). Exploring food processing in natural science education: Practical applications and pedagogical techniques. *Natural and Engineering Sciences*, 9(2), 359-375. <https://doi.org/10.28978/nesciences.1574453>
34. Mitra, A., Ammu, V., Chowdhury, R., Kumar, P. and Glory, E., 2024, August. An Adaptive Cloud and Internet of Things-Based Disease Detection Approach for Secure Healthcare system. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-7). IEEE.

## FINANCING

The authors did not receive financing for the development of this research.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## AUTHORSHIP CONTRIBUTION

*Conceptualization:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Data curation:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Formal analysis:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Research:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Methodology:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Project management:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Resources:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Software:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Supervision:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Validation:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Display:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.

*Drafting - original draft:* Barnokhon Badridinova, Kamola Azimova, Gulnoza Iskandarova, Gulruh Majidova, Xasan Abdullaev, Muso Urinov, Farida Tokhirova.