Thyroid Disease Prediction based on Feature Selection and Machine Learning

Conference Paper · December 2022

DOI: 10.1109/ICCIT57492.2022.10054746

3 authors:

Zahrul Jannat
Khulna University of Engineering and Technology

Md. Shymon Islam
Khulna University

Mst Kamrun Naher Chumki
Chemnitz University of Technology

2022 25th International Conference on Computer and Information Technology (ICCIT)
17-19 December 2022, Cox's Bazar, Bangladesh

Thyroid Disease Prediction based on Feature Selection and Machine Learning

Zahrul Jannat Peya[1,a], Md. Shymon Islam[1], Mst. Kamrun Naher Chumki[1]
[1]*Department of Computer Science and Engineering*
North Western University, Khulna-9100, Bangladesh
[a]*Department of Computer Science and Engineering*
Khulna University of Engineering and Technology, Khulna-9203, Bangladesh
Email- jannat.kuet@gmail.com, shymum1702@cseku.ac.bd, kamrunnaherchumki001@gmail.com

***Abstract*— Thyroid illness is a medical disorder in which the thyroid gland fails to produce enough hormones. Males, females, babies, teenagers, and the elderly are all susceptible to thyroid illness. It could be present from birth (hypothyroidism), or it could develop as you become older (often after menopause in women). People with thyroid diseases suffer from various problems like gaining weight, forgetfulness, anxiety, losing weight, fatigue, sleeping disorder, etc. So, diagnosing thyroid diseases is a vital issue as the diseases can be cured through proper and timely diagnosis. Recently machine learning techniques are used for diagnosing thyroid diseases. The feature selection approach was used to eliminate certain irrelevant characteristics from the thyroid dataset (from the UCI machine learning repository) and to select optimal features. The dataset has three target classes named normal, hypothyroid, and hyperthyroid. The subjects were classified through seven different machine-learning algorithms. Random Forest classifier achieves the highest accuracy 99.58% which is better than the existing state-of-the-art methods.**

Keywords*— *thyroid disease, feature selection, optimal feature, dimensionality reduction, naive bayes, random forest, machine learning

I. INTRODUCTION

Thyroid disease is a medical condition in which the thyroid gland's ability to function is harmed. The thyroid gland is an endocrine organ located in the front of the neck that produces thyroid hormones that travel through the bloodstream to help regulate a variety of other organs [1] [2]. The thyroid gland regulates metabolic activity in the human body by producing and maintaining hormone levels. Metabolism is the process by which the food we ingest is converted into energy. This energy is utilized by several of our body's systems to keep them functioning properly.

The thyroid regulates our metabolism with the help of a few hormones. Thyroxin, or T4, is a hormone with four iodine atoms. The T3 hormone, also known as triiodothyronine, is made up of three iodine atoms. The thyroid produces these two hormones, which tell the body's cells how much energy to use. When the thyroid is functioning properly, it produces the appropriate amounts of hormones to keep the body's metabolism running smoothly. Thyroid Stimulating Hormone (TSH) is another important hormone generated by the brain's pituitary gland. There are many types of thyroid diseases. When the body produces too much thyroid hormone, hyperthyroidism develops. Insufficient hormone production causes hypothyroidism. Even though the side effects can be unpleasant or stressful, most thyroid problems can be efficiently managed if they are diagnosed and treated properly [2] [3].

Thyroid problems have recently become increasingly prevalent. Thyroid disease affects 50 million people in Bangladesh, with 30 million of them unaware of their condition [4]. Around 10% of Bangladeshis are thought to have clinically apparent thyroid problems. Sub-clinical hypothyroidism and hyperthyroidism have recently been added to the list of thyroid disorders, bringing the overall number of people with thyroid problems to 20%. So proper diagnosis of thyroid disease is a major concern in today's world.

Machine Learning (ML) methods analyze data automatically and build analytical models [5] [6]. So ML methods can be used for thyroid disease diagnosis. Feature selection is essential in ML to improve efficiency. Feature selection is a method of selecting qualities based on their significance. It means that the most important elements of the machine learning model have been picked for prediction. It is utilized to get rid of irrelevant characteristics that have nothing to do with the model. There are several feature extraction techniques, of them chi-square test under filter method has been used in this study for optimal feature selection and classification is performed through the ML algorithms namely Naive Bayes (NB), Logistic Regression (LR), Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (CART) and Random Forest (RF). The contributions of the proposed work are as follows:

1) Apply the chi-square test for feature selection to eliminate certain irrelevant characteristics from the thyroid dataset and so perform the dimensionality reduction to remove complexity.
2) Apply different machine learning algorithms and achieves better accuracy than the existing state-of-the-art methods.

In this article, so far, we have demonstrated the basic ideas of Thyroid disease diagnosis in section I. Section II is for literature review, section III demonstrates our proposed method for optimal feature selection and ML-based Thyroid disease diagnosis, section IV describes the experimental results and analysis, and section V concludes the work.

II. LITERATURE REVIEW

Machine Learning methods have been used to diagnose thyroid disease in several studies. The following are some of the studies:

Tyagi et al. [7] established the Interactive Thyroid Disease Prediction System, by Using Decision Trees, Artificial Neural Networks, K-NN, and Support Vector Machine algorithms. The task has been divided into two parts as well as the thyroid illness records have been gathered from the UCI machine learning resource. The algorithm's accuracy was 75.76%, 97.05%, 98.62%, and 99.63% respectively.

M. Shyamala et al. [8] analyzed Thyroid Disease Prediction by Machine Learning Technique from Healthcare Communities, reviewed thyroid illnesses using a sample group, both unstructured and structured, and employed the FR-Growth and Decision Tree methods. To solve the problem of inadequate data, it employs a latent component model to rebuild missing data. The dataset was obtained from the UC Irvine Machine Learning Database. The FP Growth method they recommend has a 98.8% computation accuracy.

Dr. D. Jamkhandikar et al. [9] presented Thyroid Disease Prediction Using Feature Selection And Machine Learning Classifiers, using three algorithms and a feature selection technique. Applying categorization techniques from data mining, thyroid problems have been identified. SVM, Naive Bayes, and KNN all have accuracy rates of 85%, 82%, and 83% from the given set of data, respectively.

S. Razia, et al. [10] surveyed A Comparative study of machine learning algorithms on thyroid disease prediction, this work analyzes the use of a Support Vector Machine (SVM), Multiple Linear Regression, Naive Bayes, and Decision Trees in the diagnosis of a thyroid issue. The UCI machine learning database's thyroid disease dataset was applied to this. Decision Trees had a 99.23% accuracy rate.

H. Viral et al. [11], recommended Thyroid Prediction System using Machine Learning Techniques, users compare favorably the two machine learning techniques, PCA and LDA, in this study. The dataset from the UCI repository was used to train the algorithms, which were subsequently evaluated using the same dataset. They used classification models such as Random Forest, Naive Bayes, Decision Tree, K-nearest Neighbor, SVM, and Logistic Regression.

F. Dr. G. Rasitha Banu et al. [12], studied Predicting Thyroid Disease using Linear Discriminant Analysis (LDA) Data Mining Technique, and the dataset used to conduct the study on hypothyroid is taken from the UCI repository. For that experiment, they used a K-fold cross-validation approach. To improve accuracy, an experimental investigation is conducted utilizing linear discriminant analysis. The accuracy of the LDA algorithm is 99.62% while cross-validation with k = 6.

III. FEATURE SELECTION AND ML-BASED THYROID DISEASES PREDICTION

A. System Workflow

Thyroid disease is one of the most common and fast-progressive endocrine disorders. As a result, predicting diseases of this nature is a crucial endeavor. At first, we collected thyroid data of 21 features from the UCI machine learning repository. Then only 14 important features were selected through the chi-square test of the filtering method. Fig. 1 depicts the proposed methodology. Finally, classification was performed through NB, LR, ANN, KNN, SVM, CART, and RF of ML.

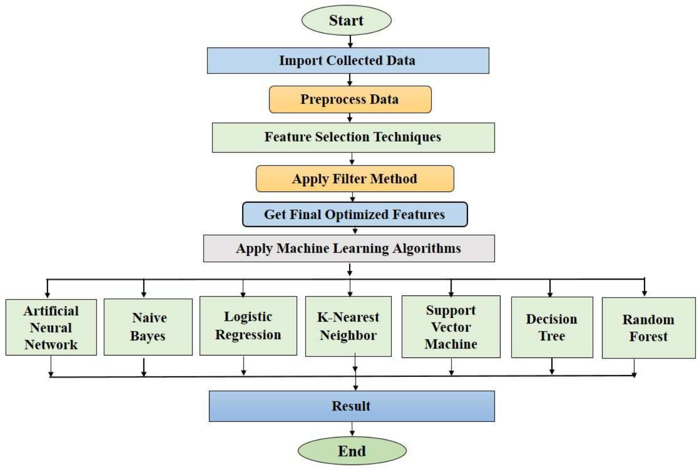

Fig. 1. Work Flow Diagram of Our Proposed Thyroid Diseases Prediction System

B. Dataset

The dataset for this study was derived from the UC Irvin information discovery in databases thyroid data collection of machine learning repository. The information was gathered using a pre-designed questionnaire. With the guidance of a thyroid medical specialist, we finalized the attributes in our dataset. The records were developed in Microsoft Excel format. The dataset includes 7200 instances, each of which has 21 descriptive attributes and a multi-class target attribute. The descriptive 21 attributes and their respective datatypes are shown in TABLE I. The dataset includes eighteen pathological and three serological attributes. There are three potential values for the multi-class target attribute: normal, hyperthyroidism, and hypothyroidism.

C. Data Preprocessing

In the initial stages of data preprocessing, we looked for any missing values in our dataset, but none were present. Then, when we looked for duplicate values, we discovered 71 of them. We subsequently discard those 71 values because they did not even contribute to algorithm training and instead increase overhead and processing time. Later, to normalize the input dataset's functional range, we used the Standard Scaler method, a crucial approach that is typically used as a preprocessing step before many machine learning models to get better accuracy.

D. *Feature Selection*

Feature selection is the process of minimizing the number of input variables when creating a predictive model. Machine learning feature selection approaches may be classified into two groups. Supervised techniques may be applied to labeled data and are used to discover relevant features in supervised models such as classification and regression. Unlabeled data can be analyzed using Unsupervised Techniques. Fig. 2 shows the feature extraction methods. Of the methods, the filter method focuses on the inherent characteristics of features as evaluated by univariate statistics. For this reason, this method has been used in this study. There are various types of filter methods for feature selection such as the chi-square test, fisher's score, information gain etc. The Chi-square test method has been used in this study for selecting features.

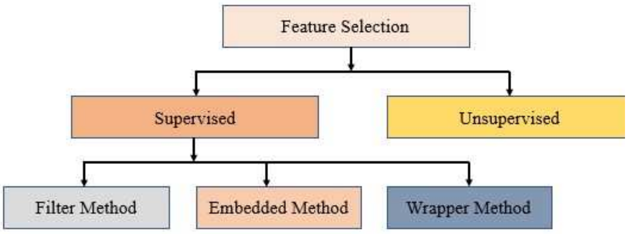

Fig. 2. Feature Selection Methods

1) Chi-square test: Data are divided into mutually exclusive groups. If the null hypothesis that no differences exist between the classes in the population is valid, the test statistic produced from the data has a $\chi 2$ frequency distribution. If the null hypothesis is valid, the test's purpose is to determine how likely the frequency components are. When the data is independent, test statistics with a $\chi 2$ distribution are generated. A chi-square test for independence evaluates two parameters in a contingency table to see if they are connected. It investigates if categorical variable distributions vary from one another in a larger sense [13]. The chi-square statistic used in the chi-square test has the following formula:

$$\chi_d{}^2 = \sum \frac{(A_x - p_x)^2}{p_x} \quad (1)$$

The subscript *d* denotes the degrees of freedom. *A* is the actual value, whereas *P* is the anticipated value. It is unlikely that you shall need to manually compute a crucial chi-square number using this algorithm. Every data item in your data collection must be calculated, as indicated by the summation sign. As you can think, the calculations may be time-consuming and tiresome [13]. The degree of freedom is a numerical value that influences the form of our distribution. It is computed as follows:

$$df = (i - 1)(j - 1) \quad (2)$$

Here, *i* is the number of rows in our contingency table and *j* is the number of columns.

E. *Train Test Splitting Methods*

In this study, we have used two different splitting methods: Holdout and k-fold cross-validation with k = 10. We have taken 80% of the data for training and 20% of data for testing in this study for the holdout method, resulting in the high accuracy of all algorithms we tested.

TABLE I. CHI-SQUARE TEST SCORE VALUE

SI	Feature Name	Datatype	Score Value	Feature Selected
1	TSH	Numeric	299.679587	**Yes**
2	On thyroxine	Numeric	52.895096	**Yes**
3	Query hypothyroid	Numeric	48.115989	**Yes**
4	Sex	Numeric	15.812910	**Yes**
5	FIT	Numeric	10.967446	**Yes**
6	Psych	Numeric	10.717321	**Yes**
7	TT4	Numeric	10.208768	**Yes**
8	Thyroid surgery	Numeric	6.504785	**Yes**
9	Pregnant	Numeric	6.315694	**Yes**
10	Sick	Numeric	4.850825	**Yes**
11	Goitre	Numeric	4.777255	**Yes**
12	I131 treatment	Numeric	3.681488	**Yes**
13	Query on thyroxine	Numeric	2.784602	**Yes**
14	Lithium	Numeric	2.488061	**Yes**
15	On antithyroid medication	Numeric	1.298478	No
16	Query hyperthyroid	Numeric	1.171358	No
17	T3	Numeric	1.076758	No
18	Age	Numeric	0.857888	No
19	Tumor	Numeric	0.273847	No
20	Hypopituitarism	Numeric	0.080970	No
21	T4U	Numeric	0.049940	No

F. *Machine Learning Algorithms*

1) Naive Bayes (NB): In the Naive Bayes classifier the existence of one feature in a class is assumed to be independent of the presence of any other feature. Even if these characteristics are connected, a Naive Bayes classifier would analyze each of them separately when determining the likelihood of a specific outcome. A Naive Bayesian model is simple to construct and can be used to analyze large datasets [14].

2) Logistic Regression (LR): Logistic Regression (LR) is used to estimate discrete values (typically binary values like 0/1) from a set of independent variables. By fitting data to a logit function, it aids in predicting the probability of an event. It's sometimes referred to as logit regression.

3) Artificial Neural Network (ANN): To grasp the concept of an artificial neural network algorithm, we must first grasp the components of a neural network. An ANN is composed of three basic layers (Input, Hidden, Output) of artificial neurons known as units. ANN are two types: Feed Forward Artificial Neural Networks and Feedback Artificial Neural Networks depending on the layer structure.

4) K-Nearest Neighbor (KNN): KNN is a straightforward algorithm that saves all existing

examples and classifies any new ones based on the votes of its k neighbors. The case is then placed in the class that shares the most similarities with it. This measurement is carried out via a distance function [14].

5) Support Vector Machine (SVM): SVM is a structural risk minimization based supervised binary classifier [15]. SVM creates a hyperplane or decision boundary to segregate n-dimensional space into classes.

6) Decision Tree (CART): A decision Tree is a supervised machine learning algorithm that creates a tree-shaped structure where the internal nodes represent dataset features, branches demonstrate decision rules and the leaf nodes define the outcomes.

7) Random Forest (RF): A Random Forest is a collection of decision trees. Each tree is classed, and the tree votes for that class, to classify a new item based on its attributes. The classification with the highest votes is chosen by the forest.

G. Evaluate Results on Different Metrics

Applying different machine learning techniques, we have measured the performance of our proposed method on different metrics such as the area under curve value, mean Squared Error value, confusion matrix, precision, recall, F1-score, accuracy, ROC curve, etc. These measurements are explicitly shown in section IV.

IV. Experimental Results and Analysis

A Chi-square test was performed on the dataset to remove irrelevant characteristics. Through the test 14 important features were selected for classifying the thyroid data. TABLE II shows the no. of selected features, Area Under Curve (AUC) Value, and Mean Squared Error (MSE) value and accuracy respectively when the holdout method is used to split the dataset into train and test sets. From the table, it is observed that the RF classifier achieved the highest accuracy of 99.5792% of others and the lowest mean squared error value of 0.01. It is also observed that the LR algorithm got the highest area under the curve value of 0.8871. So, as a whole, it can be concluded from TABLE II that the RF algorithm outperforms other algorithms when the holdout method is used. We have applied another splitting method for the dataset, namely, k-fold cross-validation with $k = 10$. Applying 10-fold cross-validation on the dataset, the similar measures, no. of selected features, an area under curve value, mean squared error value, and accuracy are being determined and these results are shown in TABLE III. From the table, it is noticed that the RF classifier achieved the highest accuracy as well as the highest area under curve value, 99.5791% and 0.9997 respectively. It is also observed that the CART algorithm got the lowest mean squared error value 0.0432. So, as a whole, it can be concluded that the Random Forest algorithm works better than other algorithms in different measures when 10-fold cross-validation is used.

We also measured performance through other measures such as precision, recall, and F1-score, and then calculated the macro average and the weighted average for three different classes normal, hyperthyroidism, and hypothyroidism. TABLE IV depicts the results of seven different classifiers on two different selection criteria. Applying feature selection along with the holdout method, for the normal Thyroid class, LR, KNN, and SVM classifiers jointly achieved the highest precision, CART got the best recall value and RF attained the best F1-score among others.

TABLE II. COMPARISON OF DIFFERENT CLASSIFIERS ON DIFFERENT MEASURES USING THE HOLDOUT METHOD

Algorithm	**Selected Features**	**AUC Value**	**MSE Value**	**Accuracy (%)**
NB	14	0.5835	0.13	93.4081
LR	14	**0.8871**	0.11	93.9691
ANN	14	0.4666	0.07	94.6704
KNN	14	0.5244	0.04	96.7741
SVM	14	0.4337	0.11	94.0392
CART	14	0.4738	0.02	99.0182
RF	14	0.5781	**0.01**	**99.5792**

TABLE III. COMPARISON OF DIFFERENT CLASSIFIERS ON DIFFERENT MEASURES USING A 10-FOLD CROSS-VALIDATION METHOD

Algorithm	**Selected Features**	**AUC Value**	**MSE Value**	**Accuracy (%)**
NB	14	0.9252	0.1595	92.2847
LR	14	0.9776	0.1414	92.7231
ANN	14	0.9636	0.1307	95.9495
KNN	14	0.6956	0.2254	96.4055
SVM	14	0.9947	0.1254	92.8108
CART	14	0.9944	**0.0432**	99.4915
RF	14	**0.9997**	0.0495	**99.5791**

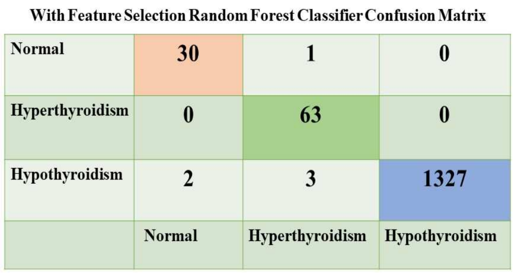

Fig. 3. Confusion Matrix of Random Forest Classifier

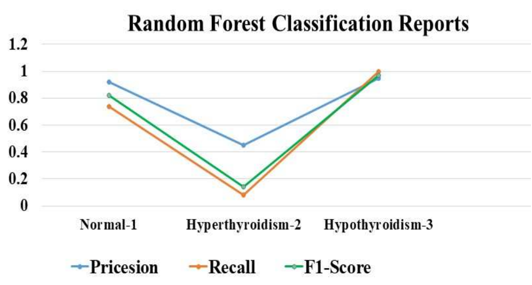

Fig. 4. Classification results of Random Forest Classifier

TABLE IV. RESULTS ANALYSIS WITH VARIOUS MEASUREMENTS

Selection Criteria	**Feature Selection along with Holdout Method**					**Feature Selection along with 10-Fold Cross Validation Method**			
Algorithm	Predicted Class	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
Naive Bayes	Normal-1	0.00	0.00	0.00	31	0.81	0.84	0.82	31
	Hyperthyroidism-2	0.00	0.00	0.00	63	0.76	0.32	0.45	63
	Hypothyroidism-3	0.93	**1.00**	0.97	1332	0.96	0.99	0.98	1332
	Accuracy			0.93	1426			0.92	1426
	Macro average	0.31	0.33	0.32	1426	0.84	0.72	0.75	1426
	Weighted average	0.87	0.93	0.90	1426	0.95	0.96	0.95	1426
Logistic Regression	Normal-1	**1.00**	0.26	0.41	31	0.82	0.77	0.79	31
	Hyperthyroidism-2	0.00	0.00	0.00	63	0.78	0.39	0.52	63
	Hypothyroidism-3	0.94	**1.00**	0.97	1332	0.97	0.99	0.98	1332
	Accuracy			0.94	1426			0.93	1426
	Macro average	0.65	0.42	0.46	1426	0.86	0.72	0.76	1426
	Weighted average	0.90	0.94	0.91	1426	0.95	0.96	0.95	1426
Artificial Neural Network	Normal-1	0.92	0.74	0.82	31	0.83	0.84	0.84	31
	Hyperthyroidism-2	0.33	0.06	0.11	63	0.81	0.66	0.73	63
	Hypothyroidism-3	0.95	0.99	0.97	1332	0.98	0.99	0.99	1332
	Accuracy			0.95	1426			0.96	1426
	Macro average	0.74	0.60	0.63	1426	0.87	0.83	0.89	1426
	Weighted average	0.93	0.95	0.93	1426	0.97	0.97	0.97	1426
K-Nearest Neighbor	Normal-1	**1.00**	0.77	0.87	31	0.84	0.66	0.74	31
	Hyperthyroidism-2	0.92	0.38	0.54	63	0.30	0.25	0.27	63
	Hypothyroidism-3	0.97	**1.00**	0.98	1332	0.59	0.95	0.97	1332
	Accuracy			0.97	1426			0.96	1426
	Macro average	0.96	0.72	0.80	1426	0.58	0.62	0.73	1426
	Weighted average	0.97	0.97	0.96	1426	0.92	0.92	0.92	1426
Support Vector Machine	Normal-1	**1.00**	0.29	0.45	31	0.87	0.65	0.74	31
	Hyperthyroidism-2	0.00	0.00	0.00	63	0.75	0.61	0.68	63
	Hypothyroidism-3	0.94	**1.00**	0.97	1332	0.98	0.99	0.99	1332
	Accuracy			0.94	1426			0.93	1426
	Macro average	0.65	0.43	0.47	1426	0.87	0.75	0.80	1426
	Weighted average	0.90	0.94	0.91	1426	0.96	0.97	0.96	1426
Decision Tree (CART)	Normal-1	0.89	**1.00**	0.94	31	**0.94**	**0.99**	**0.96**	31
	Hyperthyroidism-2	0.86	**1.00**	0.93	63	**0.99**	**0.99**	**0.99**	63
	Hypothyroidism-3	**1.00**	0.99	0.99	1332	**0.99**	**0.99**	**0.99**	1332
	Accuracy			0.99	1426			0.99	1426
	Macro average	0.92	**1.00**	0.95	1426	0.97	0.99	0.98	1426
	Weighted average	0.99	0.99	0.99	1426	0.99	0.99	0.99	1426
Random Forest	Normal-1	0.94	0.97	0.95	31	0.93	**0.99**	**0.96**	31
	Hyperthyroidism-2	**0.94**	**1.00**	0.97	63	0.98	**0.99**	0.98	63
	Hypothyroidism-3	**1.00**	**1.00**	**1.00**	1332	**0.99**	**0.99**	**0.99**	1332
	Accuracy			**1.00**	1426			0.99	1426
	Macro average	0.96	0.99	0.97	1426	0.97	0.99	0.98	1426
	Weighted average	**1.00**	**1.00**	**1.00**	1426	0.99	0.99	0.99	1426

As a whole, it can be concluded from TABLE IV that the Random Forest classifier works better than other algorithms for both holdout and 10-fold cross-validation methods. Applying feature selection along with 10-fold cross-validation, for all the 3 target classes (normal, hyperthyroidism, and hypothyroidism), CART and RF outperforms all other classifiers. Again considering the hyperthyroidism class, RF obtained the highest precision, recall as well as F1-score singly. For the hypothyroidism class, CART and RF got the highest precision, while NB, LR, KNN, SVM, and RF jointly achieved the best recall value, and RF attained the best F1 score.

TABLE V. ACCURACY OVERVIEW OF THE PROPOSED METHOD ON DIFFERENT CLASSIFIERS

Algorithm	**Accuracy without Feature Selection (%)**	**Accuracy with Feature Selection (Holdout) (%)**	**Accuracy with Feature Selection (10-fold) (%)**
NB	93.4081	93.4081	92.2847
LR	93.9691	93.9691	92.7231
ANN	96.0028	94.6704	95.9495
KNN	94.3197	96.7741	96.4055
SVM	94.0024	94.0392	92.8108
CART	99.3927	99.0182	99.4915
RF	99.5091	**99.5792**	**99.5791**

TABLE VI. ACCURACY COMPARISON AMONG EXISTING WORKS AND OUR PROPOSED WORK

Author	**Year**	**Selected Features**	**Accuracy (%)**
S. Razia et al. [10]	2018	21	73.29
A. Tyagi et al. [7]	2018	8	92.87
Y. I. Mir et al. [16]	2020	20	90.04
Dr. Dayanand et al. [9]	2020	15	83.91
G. Chaubey et al. [17]	2020	5	88.54
A. Tahir et al. [18]	2022	21	94.80
S. Sankar et al. [19]	2022	21	98.59
Our Proposed Method	2022	14	**99.58**

Figure 3 and 4 describes the results of RF in a nutshell. TABLE V demonstrates the accuracy overview of the proposed method on different classifiers. When we examined the models, we found that RF predicts 99.58% which is our gained highest accuracy with 14 features. To observe the performance of the proposed model, we have compared our proposed method with several recent existing methods and the results are shown in TABLE VI briefly.

V. CONCLUSION

Thyroid disease affects a large number of people in Bangladesh. The majority of them are completely unaware that they have it. It would be extremely beneficial to them if it could be accurately predicted. This research presents the use of feature selection in thyroid patient's data which are used in the machine learning models for classification. For all the features that contribute to the feature set at a high computational cost, the volume of data is enormous. The number of features is decreased, which directly affects data reduction and significantly lowers computing costs. For this reason, this study investigates the Chi-square test method for selecting the important features from the dataset. Machine learning algorithms namely NB, LR, ANN, KNN, SVM, CART, and RF were applied to the data of selected features for performing a three-class classification problem (Normal, hypothyroidism, hyperthyroidism). The best accuracy of 99.58% was achieved through RF. A future extension of this research would be to apply this approach to enlarge datasets and for diagnosing several other diseases.

REFERENCES

[1] V. Prasad, T. S. Rao, and M. S. P. Babu, "Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms," Soft Comput., vol. 20, no. 3, pp. 1179–1189, 2016, doi: 10.1007/s00500-014-1581-5.

[2] "https://my.clevelandclinic.org/health/diseases/8541-thyroid-disease."

[3] L. D. K. E. Premawardhana and J. H. Lazarus, "Management of thyroid disorders," Postgrad. Med. J., vol. 82, no. 971, pp. 552–558, 2006, doi: 10.1136/pgmj.2006.047290.

[4] "50 million people suffer from thyroid disease in Bangladesh," Dhaka Tribune, Dhaka, 2018.

[5] B. P. Woolf, "Chapter 7 - Machine Learning," in Building Intelligent Interactive Tutors, B. P. Woolf, Ed. San Francisco: Morgan Kaufmann, 2009, pp. 221–297.

[6] M. A. Myszczynska et al., "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases," Nat. Rev. Neurol., vol. 16, no. 8, pp. 440–456, 2020, doi: 10.1038/s41582-020-0377-8.

[7] A. Tyagi, R. Mehra, and A. Saxena, "Interactive thyroid disease prediction system using machine learning technique," PDGC 2018 - 2018 5th Int. Conf. Parallel, Distrib. Grid Comput., pp. 689–693, 2018, doi: 10.1109/PDGC.2018.8745910.

[8] M. Shyamala and P. S. S. Akilashri, "Thyroid Disease Prediction by Machine Learning Technique From Healthcare Communities," Int. J. Comput. Sci. Eng., vol. 6, no. 11, 2018.

[9] D. Jamkhandikar and N. Priya, "Thyroid Disease Prediction Using Feature Selection and Machine Learning Classifiers," Int. J. Anal. Exp. modal Anal., vol. XII, no. 0886, pp. 175–180, 2020.

[10] S. Razia, P. Swathi Prathyusha, N. Vamsi Krishna, and N. Sathya Sumana, "A Comparative study of machine learning algorithms on thyroid disease prediction," Int. J. Eng. Technol., vol. 7, no. 2.8, pp. 315–319, 2018, doi: 10.14419/ijet.v7i2.8.10432.

[11] M. S. P. B. and Haria Viral, "Thyroid Prediction System using Machine Learning Techniques," Int. J. Sci. Res. Rev., vol. 7, no. 4, pp. 601–608, 2016.

[12] G. Rasitha Banu, "Predicting Thyroid Disease using Linear Discriminant Analysis(LDA) Data Mining Technique," Communications on Applied Electronics (CAE): Vol. 4, no. 12, January 2016.

[13] "Chi-Square Statistic: How to Calculate It / Distribution," www.statisticshowto.com, 2021.

[14] M. J. Islam, Q. M. J. Wu, M. Ahmadi and M. A. Sid-Ahmed, "Investigating the Performance of Naive- Bayes Classifiers and KNearest Neighbor Classifiers," no. June 2014, pp. 1541–1546, 2008, doi: 10.1109/iccit.2007.148.

[15] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Apr. 1995.

[16] Y. I. Mir, Dr. S. Mittal, "Thyroid Disease Prediction Using Hybrid Machine Learning Techniques: An Effective Framework," Int. J. of scientific & technology research, vol. 9, no. 02, February 2020, pp. 2868-2874, 2020.

[17] G. Chaubey, D. Bisen, S. Arjaria and V. Yadav, "Thyroid Disease Prediction Using Machine Learning Approaches," in The National Academy of Sciences, India, 2020.

[18] T. Alyas, M. Hamid, K. Alissa, T. Faiz, N. Tabassum and A. Ahmad "Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach" Hindawi BioMed Research International, Volume 2022, Article ID 9809932, 10 pages, https://doi.org/10.1155/2022/9809932

[19] S. Sankar, A. Potti, G. N. Chandrika, and S. Ramasubbareddy "Thyroid Disease Prediction Using XGBoost Algorithms" Journal of Mobile Multimedia, vol. 18, no. 3, https://doi.org/10.13052/jmm1550-4646.18322.