

PAPER

Novel Classification Approach for Thyroid Detection: Feature Enhanced AdaBoost Optimization with Max Voting

Deepali Bhende¹(✉),
Gopal Sakarkar^{1,2}, Ambika
Jaiswal³, Punam Khandar⁴,
Satyajit Uparkar⁴,
Lalit Agrawal⁴

¹G H Raisoni University,
Saikheda, Madhya
Pradesh, India

²Dr. Vishwanath Karad,
MIT World Peace University,
Pune, Maharashtra, India

³Sant Gadge Baba Amravati
University, Amravati,
Maharashtra, India

⁴Shri Ramdeobaba College
of Engineering and
Management, Nagpur,
Maharashtra, India

deepali.bhende.phdcs@ghru.edu.in

ABSTRACT

The need for enhanced methods in disease prediction is a significant challenge in the medical field. Current predictive models often face challenges such as limited accuracy, insufficient adaptability to diverse datasets, and inefficiencies in feature selection and model training. These limitations can hinder early diagnosis and effective management of thyroid conditions, which are vital for patient outcomes. The study introduces an innovative method for enhancing thyroid disease prediction using a machine learning study employs algorithms such as support vector machine (SVM), Naive Bayes (NB), K-nearest neighbor (KNN), logistic regression (LR), and stochastic gradient descent (SGD) in conjunction with filter, wrapper, and embedded feature selection methods across three distinct models. The study uses two thyroid datasets, one from Dew Medicare Ternity Hospital, Nagpur, and the other from the UCI thyroid repository, revealing the potential of the novel 'FeatureBoostThyro' approach for improving thyroid risk prediction across diverse datasets. The proposed method achieved accuracies of 98.10%, 97.47%, and 95.58% for the three models using the UCI dataset, and 97.42%, 98.71%, and 97.83% for the DMTH dataset. The novelty of this approach lies in its integrated pipeline that ensures the selection of the best features, systematic model training, and rigorous evaluation. This results in a robust, accurate, and reliable model that outperforms traditional approaches, making it a significant advancement in the field of disease prediction. The enhanced performance metrics, especially accuracy, highlight the potential of this method in clinical settings for early and accurate thyroid disease detection.

KEYWORDS

machine learning (ML), filter selection methods, wrapper feature selection, logistic regression (LR), support vector machine (SVM), AdaBoost, GridSearchCV

Bhende, D., Sakarkar, G., Jaiswal, A., Khandar, P., Uparkar, S., Agrawal, L. (2024). Novel Classification Approach for Thyroid Detection: Feature Enhanced AdaBoost Optimization with Max Voting. *International Journal of Online and Biomedical Engineering (ijOE)*, 20(14), pp. 71–84. <https://doi.org/10.3991/ijoe.v20i14.50623>

Article submitted 2024-06-19. Revision uploaded 2024-08-31. Final acceptance 2024-08-31.

© 2024 by the authors of this article. Published under CC-BY.

1 INTRODUCTION

Multiple endocrine glands secrete different hormones, each controlling various bodily functions. The thyroid gland is one of the most significant endocrine glands among them [1]. It secretes three hormones: tri-iodothyronine, thyroxine, and calcitonin. Low secretion of these hormones results in hypothyroidism, while excessive secretion leads to hyperthyroidism [2]. Other thyroid disorders include goiter, Hashimoto's thyroiditis, and thyroid cancer. The thyroid gland's major job is to keep the bloodstream stable by controlling metabolism [3]. Improper levels of thyroid hormones can lead to disorders that are dangerous and require medical attention [4]. Therefore, accurate diagnosis is critical and must be performed by experts. Recent advances in machine learning (ML) are transforming healthcare by analyzing vast data to predict diseases early. This study aims to find the best algorithm for predicting thyroid disorders using logistic regression (LR), support vector machine (SVM), K-nearest neighbors (KNN), Naive Bayes (NB), and stochastic gradient descent (SGD). Various feature selection methods were tested, and while initial models did not yield a clear best algorithm, an optimized ensemble model showed consistent improvement. The models were implemented in Python and tested on datasets from Dew Medicare Trinity Hospital and the UCI repository. This paper is organized as follows: In Section 2, relevant research on thyroid problems is included. Section 3 discusses the study's materials and procedures. Section 4 provides a full summary of the study's findings, and Section 5 concludes the study.

2 RELATED WORKS

Ege Savc et al. [4] argued that under-sampling might not be the optimal technique for handling small, imbalanced datasets. They noted that both the training and test datasets were imbalanced and suggested that the best approach to avoid oversampling issues is to combine oversampling with feature importance analysis to reduce the risk of overfitting. To achieve this, the researchers excluded non-important features, identifying and retaining only the important ones. They also pointed out that certain algorithms required scaling due to their structural needs. While artificial neural networks proved to be the most effective tool for diagnosing thyroid disease, both SVM and KNNs also achieved impressive results, with accuracy rates exceeding 95.00%. When properly tuned, these algorithms are highly effective for multi-class classification tasks.

Kalpna Guleria et al. [5] said that it takes a lot of time to manually analyze parameters on big databases in order to diagnose and forecast hypo- or hyper-thyroidism. This study constructed predictive models using a variety of ML based methods. These included a model artificial neural network based on deep learning, a random forest, a NB/multiclass classifier, and a decision tree. According to the performance review, the random forest or decision tree produced better outcomes, with 99.3107% accuracy being the lowest and 99.57%) the highest. The competitive accuracy of the DL-based ANN model is 93.82%. Researchers may find this study useful in determining an appropriate model for the identification and categorization of hypothyroidism.

R. Vanitha et al. [6] The main goal of this work is to clean up data on thyroid cancer from the UCI ML repository. The procedure entails using and contrasting various algorithms to eliminate redundant values, fill in missing values, and choose the best characteristics from the recently created medical dataset. The thyroid dataset was preprocessed in this work using the KNN and PCA methods to eliminate superfluous variables. The KNN, mean/median, and PCA procedures were also compared. Out of the three imputation methods, the hybrid imputes showed superior performance.

In order to produce appropriate diagnosis results, regression and classification algorithms will be merged in future cases.

Kwang-Sig Lee et al. [7] presented this study reviewing recent advances in ML for early diagnosis of thyroid disease. Several methods include RF and gradient boosting for numerical, genomic, and random data. Accuracy ranged from 66.80 % to 90.01%. This review shows that ML offers a non-invasive decision support system for early diagnosis, including the thyroid gland.

Shiuh Tong Lim et al. [8] aimed to demonstrate the performance of the Featurewiz library and determine a remarkable model for predicting thyroid diseases among several ML models, for example, a decision tree. A dataset of Australian thyroid patients was used to develop the ML models, and the models were then built in two ways: without feature selection and with feature selection. The results of these two operations were similar to tree-based operations models. Initially, without feature selection, the basic XGBoost model is 99.23% accurate, while random forest is second with 98.79% accuracy.

Ibrahim et al. [9] stated that ML algorithms have become very important in the healthcare sector, especially in disease diagnosis. Many companies use these technologies to predict diseases early and improve medical diagnosis. The motivation of this paper was to provide an overview of ML algorithms used for detection and prediction. Several ML algorithms, such as NB, LR, SVM, KNN, artificial cluster, decision tree, and random forest, were reviewed for detection of diseases during the last three years. Comparison of these algorithms, evaluation processes, and the results obtained.

Rajasekhar Chaganti et al. [10] explained that the existing methods of model optimization are less investigated. Thus, this study presents work that explores feature selection for ML and deep learning models. Extra tree classifiers and ML-based feature selection was adopted. The proposed work predicts Hashimoto's thyroiditis. The best result was reflected by an extra tree classifier with 99% accuracy when a random forest classifier was used.

Alyaa Dawood et al. [11] proposed this study to identify the hormonal activity of the thyroid gland in hyperthyroidism and hypothyroidism. This was accomplished using several ML algorithms and classifiers, namely LR, decision trees, SVM, etc. Python simulates all algorithms in the Anaconda environment using the Spyder platform by comparing and choosing the most accurate. DT and RF algorithms gave the best results, reaching 99.33% and 99.73% accuracy, respectively.

Lerina Aversanoa et al. [12] predicted the LT4 treatment for hyperthyroid patients. The dataset was built using data from AOU Federico. Several ML algorithms were applied with 10 different classifiers. Out of all the classifiers, the extra-tree classifier reached an accuracy of 84%.

S. Sankar et al. [13] explained that detecting any disease early could help treat and prevent the worst health conditions. ML plays a crucial role in detecting diseases in their early stages. The dataset was taken from the UC Irvine Information database. The researchers used the XGBoost algorithm for accurate prediction of the thyroid. The efficiency of XGBoost was compared to the decision tree, LR, and KNN methods, and then all four were analyzed. The conclusion was that the XGBoost algorithm's performance accuracy increased by 2% compared to K-nearest neighbor.

3 MATERIALS AND METHODS

This approach introduces a novel method for enhancing classification performance by systematically integrating feature selection and ensemble learning techniques within a pipeline framework.

3.1 Support vector machine

One of the algorithms for supervised ML is SVM [14]. The approach locates the ideal hyperplane in an N-dimensional space given a feature space. In an attempt to maximize the distance between the nearest points in each class, this hyperplane divides the data points into several classes. The total number of features in the data collection determines the hyperplane's dimension. A hyperplane is referred to as a hard margin when data points can be separated perfectly. A soft margin is present when there is some degree of separability between the data points.

3.2 Logistic regression

Logistic regression, the categorical dependent variable, is predicted by the simple and understandable supervised learning approach, LR [15]. It provides probabilistic values that range from zero to one. It is applied to the resolution of classification issues. It matches the logistic function with a "S" form, which forecasts two maximum values: zero and one. This reveals whether or not a gland, for example, is cancerous. The threshold value notion is utilized in LR. The chance of either zero or one is defined by the threshold value. If the value is higher than the threshold, it usually tends to be one. If the value is less than the threshold, it usually tends to be zero. An equation between x and y represents the logistic function. The formula represents y as a sigmoid.

$$F(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

3.3 K-nearest neighbor

K-nearest neighbor is a supervised learning algorithm. It relies on the basic idea that similar data values tend to have similar data labels. It assigns the new data point to the majority set within its neighbors. As a regression algorithm, it predicts based on the average values closest to the new point [16].

3.4 Naïve bayes

Naive bayes based on the Bayes theorem, the NB algorithm facilitates supervised learning. It is applied to difficulties involving classification. This method assumes that all features are independent, so it is not used to learn the relationship between features. Bayes' theorem is used to build a learner ML model from an available set of features that, given a fresh collection of attributes, forecasts the likelihood that a target variable will belong to a certain class [17].

3.5 Stochastic gradient descent

It is used to optimize ML models. It uses a single random sample or batch of samples. Thus, it is less expensive as it processes fewer data points. The noise is high due to frequent updates with single or few samples. It requires less memory as it possesses fewer data points at a time [18]. It is less sensitive to initial parameter values due to frequent updates.

3.6 AdaBoost

Adaboost ensemble model: It is an ensemble learning method. It is adaptive boosting. It iteratively trains the weak classifiers. With the training dataset, iteratively trains the weak classifier. It gives more weight to the data points that are misclassified [19]. In addition to boosting weak learners, getting better accuracy by fine-tuning hyperparameters is possible.

3.7 GridSearchCV

Grid search optimization: In ML, models' hyperparameters are fine-tuned using different techniques. grid search cross-validation is a powerful technique for tuning the hyperparameters [19]. Hyperparameters guide the learning process. Prior to the model being trained, these parameters are set. It takes time and effort to set the hyperparameter settings. The procedure is automated by Grid Search CV. Every possible combination of hyperparameters is subject to a methodical evaluation of the model's performance. It aids in determining the ideal combination of hyperparameters. This automatically improves the performance of the model on unseen data. The key steps of this approach are as shown in the following architecture:

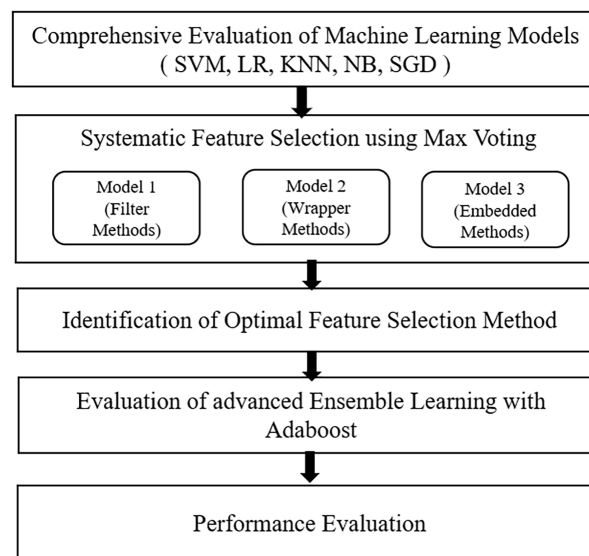


Fig. 1. System architecture

As shown in Figure 1 system architecture initially, there is assessment of the performance of various ML models, including KNN, LR, SVM, NB and SGD, on the thyroid dataset. This evaluation is performed without any feature selection to establish baseline metrics such as accuracy. In the second phase, a variety of feature selection methods are, including information gain, chi-square, gain ratio, relief, feature selection, backward selection, recursive feature elimination, and LASSO. Each method is applied to identify the most relevant features for each ML model. The overall approach is divided into three models: model 1, model 2, and model 3.

Model 1: Filter selection methods: Filter-based feature selection techniques, such as information gain, Gini index, chi square, and relief, were employed in the first model. Using these techniques, features were chosen according to how statistically significant they were for the target variable.

Model 2: Wrapper selection methods: The second model employed wrapper-based feature selection methods, including backward selection, forward selection, and recursive feature elimination. These methods involved iteratively selecting features based on their performance in training ML models.

Model 3: Embedded feature selection method: The third model used an embedded feature selection method, specifically least absolute shrinkage and selection operator. LASSO performs feature selection as part of the model training process by imposing a penalty on the absolute size of the regression coefficients.

Model 4: Novel 'FeatureBoostThyro' model: The fourth model is the novel 'FeatureBoostThyro' approach. This model combined the best feature selection method identified from the previous models with an AdaBoost ensemble model. A Random Forest Classifier was used as the base estimator, and GridSearchCV was employed to optimize the n_estimators and learning_rate parameters.

Each ML model and feature selection method is evaluated, with the most effective method chosen based on highest accuracy across models. This method is then used to transform the dataset for final model training. A random forest classifier is used in an AdaBoost ensemble, fine-tuned with GridSearchCV for optimal parameters. The best AdaBoost model is tested for accuracy. The 'FeatureBoostThyro' approach was assessed on two datasets: Dew Medicare Ternity Hospital's clinical data and the UCI Thyroid Disease Repository.

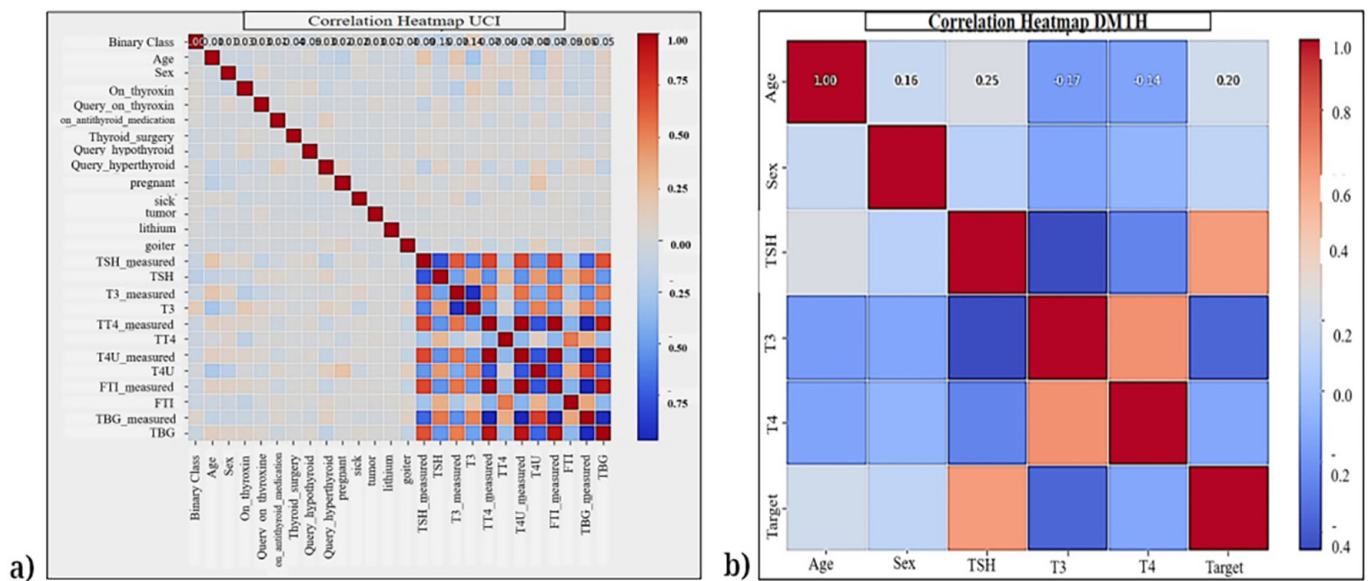


Fig. 2. (a) Correlation heatmap UCI and (b) Correlation heatmap DMTH

Both datasets were preprocessed with imputation, normalization, and encoding. Figure 2a and b show the correlation heat maps for the UCI and DMTH datasets. The UCI heat map highlights strong negative correlations between TSH and TBG and positive correlations between TT4 and T4U. The DMTH heat map reveals weak to moderate correlations among attributes, such as weak links between age, sex, and thyroid measures, and more complex relationships among TSH, T3, and T4.

4 RESULT AND DISCUSSION

Findings of the experiments are summarized as in Table 1 as performance of a secondary dataset named UCI thyroid dataset and Table 2 as performance of a primary dataset named DMTH dataset. As a key finding, the results show that different ML algorithms respond differently to various feature selection methods, with some methods improving performance, others having no effect, and some even degrading performance.

4.1 Performance on secondary dataset (UCI repository)

Table 1 shows feature selection effects on ML algorithms with the UCI dataset. KNN’s performance varied slightly, while FBT consistently performed best, maintaining 98.10% accuracy across most methods. Notably, NB saw significant improvements with methods such as IG and chi-square. Conversely, some methods hurt performance, with SGD declining with chi-square and LR with several methods.

Table 1. Performance comparison of filter, wrapper and embedded feature selection method for UCI dataset

UCI	Before IG (A0)	After IG (A1)	(A1-A0) %	UCI	Before GI	After GI	(A1-A0) %
SVM	94.00	96.37	2.37	SVM	94.00	94.15	0.15
KNN	96.21	96.21	0	KNN	96.21	95.10	-1.11
LR	94.63	94.15	-0.48	LR	94.63	94.31	-0.32
NB	27.17	94.15	66.98	NB	27.17	29.23	2.06
SGD	93.68	94.15	0.47	SGD	93.68	93.36	-0.32
FBT	98.10	98.10	0	FBT	98.10	98.10	0
UCI	Before Chi Square	After Chi Square	(A1-A0) %	UCI	Before ReF	After ReF	(A1-A0) %
SVM	94.00	96.37	2.37	SVM	94.00	94.15	0.15
KNN	96.21	96.52	0.31	KNN	96.21	95.1	-1.11
LR	94.63	94.15	-0.48	LR	94.63	94.31	-0.32
NB	27.17	94.31	67.14	NB	27.17	29.23	2.06
SGD	93.68	89.89	-3.79	SGD	93.68	94.94	1.26
FBT	98.10	98.10	0.00	FBT	98.10	98.10	0.00
UCI	Before FS (A0)	After FS (A1)	(A1-A0) %	UCI	Before BE (A0)	After BE (A1)	(A1-A0) %
SVM	94.00	94.00	0.00	SVM	94.00	94.00	0.00
KNN	96.21	95.73	-0.48	KNN	96.21	95.73	-0.48
LR	94.63	94.15	-0.48	LR	94.63	94.15	-0.48
NB	27.17	94.15	66.98	NB	27.17	94.15	66.98
SGD	91.47	92.58	1.11	SGD	91.47	92.58	1.11
FBT	97.00	97.00	0.00	FBT	97.00	97.00	0.00

(Continued)

Table 1. Performance comparison of filter, wrapper and embedded feature selection method for UCI dataset (Continued)

UCI	Before RFE (A0)	After RFE (A1)	(A1-A0) %	UCI	Before LASSO (A0)	After LASSO (A1)	(A1-A0) %
SVM	94.00	96.37	2.37	SVM	94.94	95.42	0.48
KNN	96.21	96.52	0.31	KNN	94.15	94.15	0.00
LR	94.63	94.15	-0.48	LR	97.16	96.84	-0.32
NB	27.17	94.15	66.98	NB	96.84	95.42	-1.42
SGD	91.47	94.15	2.68	SGD	96.05	96.52	0.47
FBT	97.00	97.47	0.47	FBT	97.47	95.58	-1.89

The findings show that feature selection techniques can significantly affect how well ML algorithms function. NB, in particular, benefits greatly from feature selection, especially with methods such as IG and chi-square, demonstrating substantial performance improvements. In contrast, other ML algorithms show more varied responses, with some showing improvements, some remaining unaffected, and others experiencing performance drops.

4.2 Performance on primary dataset (DMTH dataset)

Table 2 highlights that traditional ML algorithms such as SVM, KNN, LR, and NB show little to no improvement with feature selection, while SGD benefits from several methods, and FBT consistently performs well. FBT's performance remains stable across various feature selection methods, whereas SGD shows improvements of 3.23% to 5.81% with some methods, despite a slight decrease with LASSO. LR's performance drops with certain methods, and some methods negatively affect specific algorithms.

Table 2. Performance comparison of filter, wrapper and embedded feature selection method for DMTH dataset

DMTH	Before IG (A0)	After IG (A1)	(A1-A0) %	DMTH	Before GI (A0)	After GI (A1)	(A1-A0) %
SVM	87.10	87.10	0	SVM	87.10	87.10	0
KNN	91.61	91.61	0	KNN	91.61	91.61	0
LR	81.29	81.29	0	LR	81.29	81.29	0
NB	78.71	78.71	0	NB	78.71	78.71	0
SGD	76.13	80.65	4.52	SGD	76.13	80.00	3.87
FBT	97.42	97.42	0	FBT	97.42	97.42	0
DMTH	Before Chi Square (A0)	After Chi Square (A1)	(A1-A0) %	DMTH	Before ReF	After ReF	(A1-A0) %
SVM	87.10	87.10	0	SVM	87.10	87.10	0
KNN	91.61	91.61	0	KNN	91.61	91.61	0
LR	81.29	81.29	0	LR	81.29	81.29	0
NB	78.71	78.71	0	NB	78.71	78.71	0
SGD	76.13	80.65	4.52	SGD	76.13	81.94	5.81
FBT	97.42	97.42	0	FBT	97.42	97.42	0

(Continued)

Table 2. Performance comparison of filter, wrapper and embedded feature selection method for DMTH dataset (*Continued*)

DMTH	Before FS (A0)	After FS (A1)	(A1-A0) %	DMTH	Before BE (A0)	After BE (A1)	(A1-A0) %
SVM	85.16	85.16	0	SVM	85.16	85.16	0
KNN	90.32	90.32	0	KNN	90.32	90.32	0
LR	83.87	81.94	-1.93	LR	83.87	81.94	-1.93
NB	80.65	82.58	1.93	NB	80.65	82.58	1.93
SGD	79.35	82.58	3.23	SGD	79.35	83.23	3.88
FBT	97.42	97.42	0	FBT	97.42	97.42	0
DMTH	Before RFE (A0)	After RFE (A1)	(A1-A0) %	DMTH	Before LASSO (A0)	After LASSO (A1)	(A1-A0) %
SVM	85.16	85.16	0	SVM	87.10	87.10	0
KNN	90.32	90.32	0	KNN	89.68	89.68	0
LR	83.87	83.87	0	LR	83.87	83.87	0
NB	80.65	80.65	0	NB	80.65	80.65	0
SGD	79.35	83.87	4.52	SGD	82.58	82.23	-0.35
FBT	97.42	98.71	1.29	FBT	98.71	97.83	-0.88

4.3 Performance evaluation

A variety of evaluation criteria, such as accuracy, precision, recall, f1 score, and others, are available to assess how well ML models perform [19]. Accuracy was the evaluation metric employed in this investigation. How well a classifier can estimate the total population with or without thyroid depends on its accuracy.

$$Accuracy = \frac{TN + TP}{TP + FP + TN + FN} \quad (2)$$

Where TP , TN , FP , and FN are defined as follows:

- True positive (TP): The diagnostic model yields positive test results for thyroid, and the patient has thyroid.
- False positive (FP): The diagnostic model gives a positive test result for thyroid, and the patient has no thyroid.
- True negative (TN): The diagnostic model gives negative test results for thyroid, and the patient has no thyroid.
- False negative (FN): The diagnostic model gives negative test results for thyroid, and the patient has thyroid.

Figures 3–5 demonstrate that the effectiveness of feature selection methods can vary significantly between different ML algorithms when using the DMTH dataset. However, the consistently high performance of the novel FBT method suggests that it is a reliable choice across various scenarios.

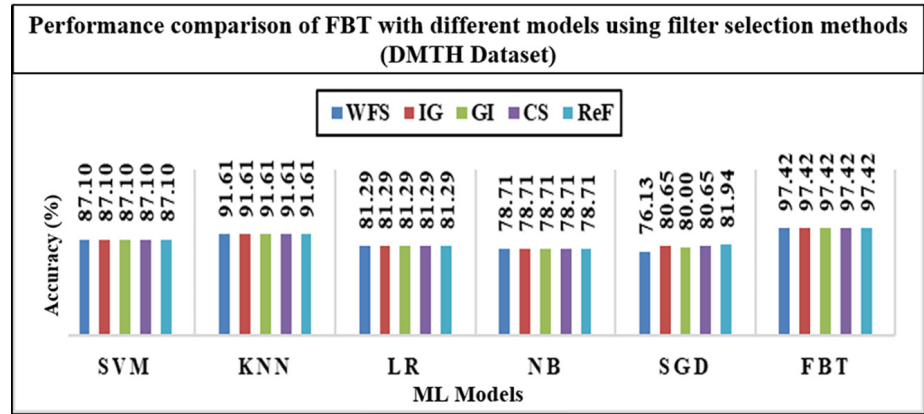


Fig. 3. Performance of model 1 (DMTH)

As shown in Figure 3 performance of model 1 varied across different filter methods, with SVM and KNN showing relatively higher accuracy.

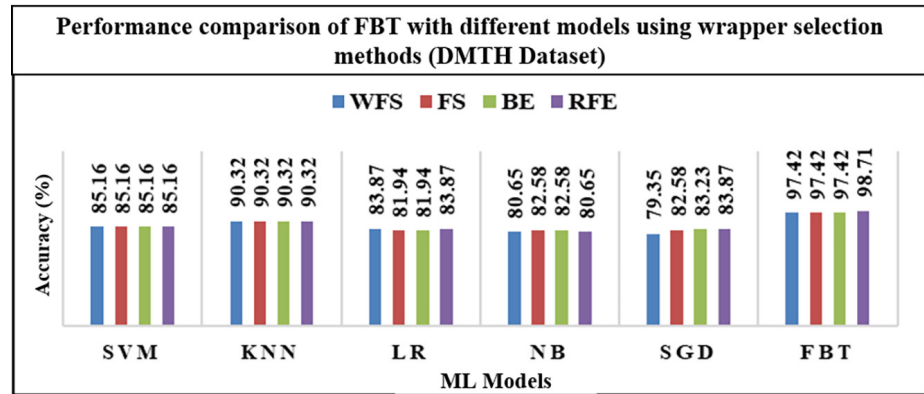


Fig. 4. Performance of model 2 (DMTH)

As shown in Figure 4, wrapper methods generally provided improved performance, with RFE often yielding the best results among this category.

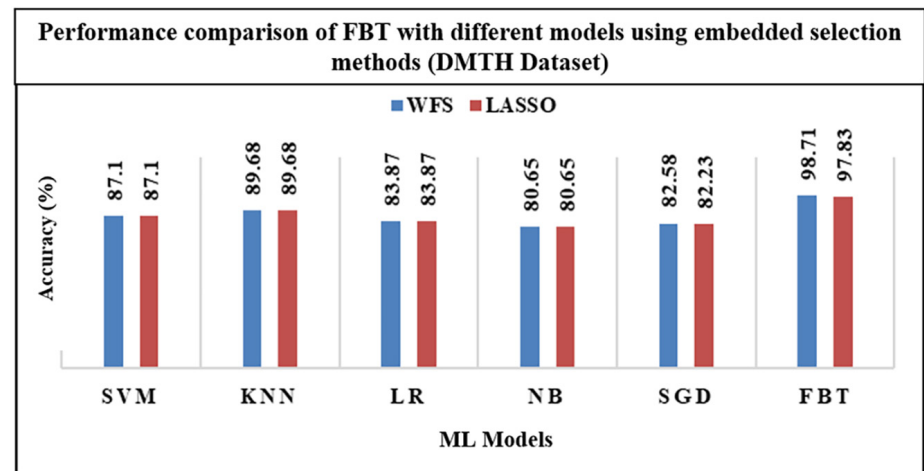


Fig. 5. Performance of model 3 (DMTH)

As shown in Figure 5, the LASSO method demonstrated competitive performance, particularly with logistic regression.

Overall, the proposed ‘FeatureBoostThyro’ model based on the AdaBoost model, optimized using GridSearchCV, demonstrated superior performance across all metrics compared to the individual ML models. The use of the best feature set from previous evaluations significantly improved the performance of the FBT model.

The performance of all three models was also evaluated using the UCI dataset. Figures 6–8 highlight the robustness and reliability of these ML algorithms when combined with different feature selection methods in terms of accuracy.

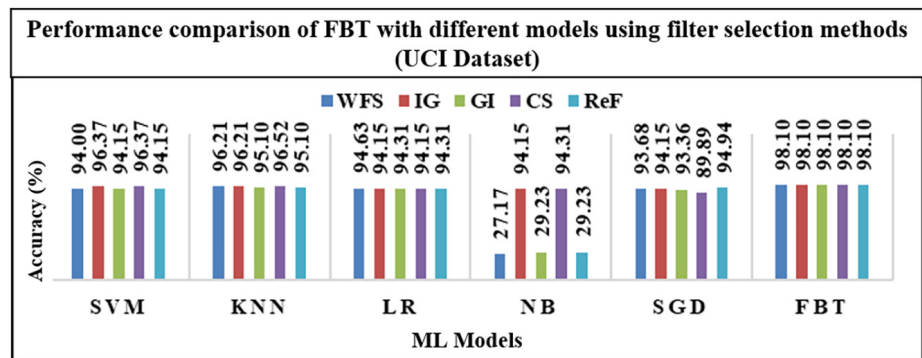


Fig. 6. Performance of model 1 (UCI)

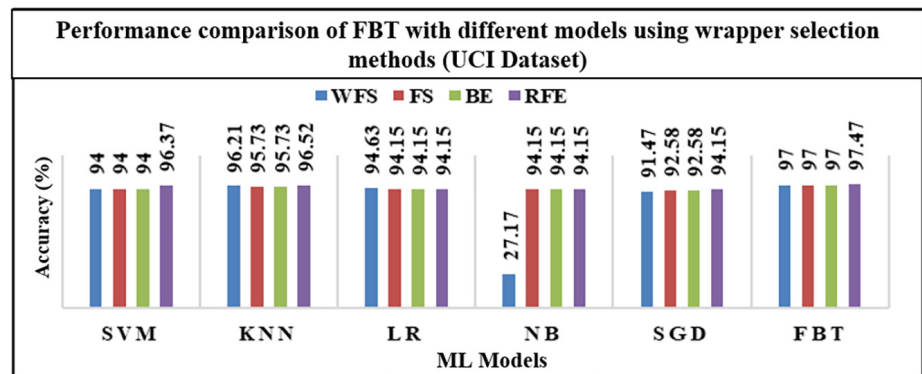


Fig. 7. Performance of model 2 (UCI)

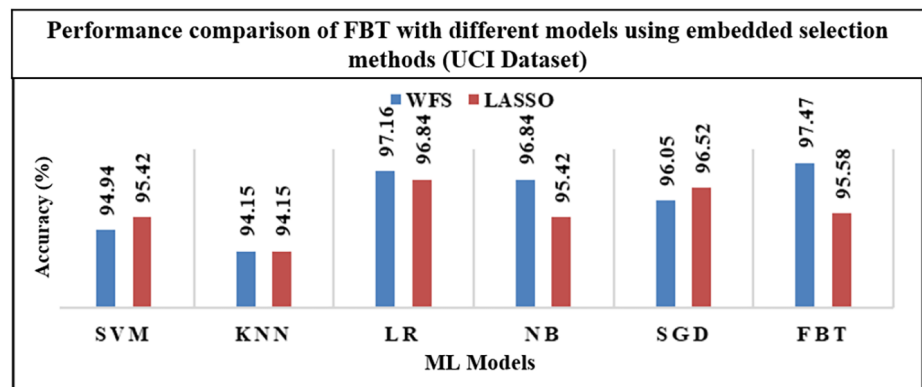


Fig. 8. Performance of model 3 (UCI)

Trends observed on the secondary dataset are:

- **Model 1 (filter methods):** SVM and KNN performed better among the individual filter methods, as shown in Figure 6.
- **Model 2 (wrapper methods):** Wrapper methods again showed improved performance, with RFE often leading as shown in Figure 7.
- **Model 3 (embedded method-LASSO):** LASSO provided robust results, particularly with LR as shown in Figure 8.

As shown in the results for all models, the proposed FBT method outperforms all other methods.

Table 3 shows the comparative analysis of the proposed FBT with state-of-the-art methods in terms of accuracy metrics.

Table 3. Performance comparison of proposed FBT with the state of art methods

Author	Algorithm	F S Method Used	Dataset	Accuracy
A. K. Chaudhuri et al. (2021) [17]	NB, SVM, LR, DT, RF, GDB Extra Tree	Information Gain, Relief F, Gain Ratio, One R	UCI	90.32%
R. Chaganti et al. (2022) [10]	RF, LR, SVM, AD, GB, LSTM, CNN	FS, BFE, BiDFE, Extra Tree, CNN	UCI	97.00%
T. Alyas et al. (2022) [1]	KNN, ANN, NB, RF		UCI	94.80%
Proposed FeatureBoostThyro (FBT) Method	AdaBoost Ensemble with GridSearchCV	Information Gain, Gain Index, Chi-Square, Relief F, FS, BFE, RFE, LASSO	UCI	98.10%
			DMTH	98.71%

5 CONCLUSIONS

In the rapidly evolving field of medical diagnostics, accurate and early detection of diseases is crucial for effective treatment and improved patient outcomes. Advances in predictive modeling and ML are playing a transformative role in enhancing diagnostic capabilities. Among these innovations, the ‘FeatureBoostThyro’ approach stands out for its promising results in thyroid disease prediction. The ‘FeatureBoostThyro’ approach, which combines comprehensive feature selection with an optimized AdaBoost ensemble model, demonstrates significant improvements in thyroid disease prediction across diverse datasets. The novelty of this approach lies in its integrated pipeline that ensures the selection of the best features, systematic model training, and rigorous evaluation. The enhanced performance metrics, especially accuracy, highlight the potential of this method in clinical settings for early and accurate thyroid disease detection. Future work could explore the application of this approach to other diseases and the integration of additional data sources to further improve prediction accuracy.

6 REFERENCES

- [1] 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. 1, 2022. <https://doi.org/10.1155/2022/9809932>
- [2] Z. J. Peya, M. S. Islam, and M. K. N. Chumki, “Thyroid disease prediction based on feature selection and machine learning,” in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, 2022, pp. 495–500. <https://doi.org/10.1109/ICCIT57492.2022.10054746>

- [3] K. Salman and E. Sonuç, "Thyroid disease classification using machine learning algorithms," *Journal of Physics: Conference Series*, vol. 1963, no. 1, p. 012140, 2021. <https://doi.org/10.1088/1742-6596/1963/1/012140>
- [4] E. Savcı and F. Nuriyeva, "Diagnosis of thyroid disease using machine learning techniques," *Journal of Modern Technology & Engineering*, vol. 7, no. 2, pp. 134–145, 2022.
- [5] 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, p. 100482, 2022. <https://doi.org/10.1016/j.measen.2022.100482>
- [6] R. Vanitha and K. Perumal, "Hybrid model of KNN and PCA to pre-processor of thyroid dataset using machine learning," *Journal of Pharmaceutical Negative Results*, vol. 13, no. 3, pp. 1091–1096, 2022. <https://doi.org/10.47750/pnr.2022.13.S03.170>
- [7] K. S. Lee and H. Park, "Machine learning on thyroid disease: A review," *Frontiers in Bioscience-Landmark*, vol. 27, no. 3, p. 101, 2022. <https://doi.org/10.31083/j.fbl2703101>
- [8] T. L. Shiuh, W. K. Khai, Y. C. Xin, and C. Y. Wai, "Prediction of thyroid disease using machine learning approaches and Featurewiz selection," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 15, no. 3, pp. 9–16, 2023. <https://doi.org/10.54554/jtec.2023.15.03.002>
- [9] I. Ibrahim and A. Abdulazeez, "The role of machine learning algorithms for diagnosing diseases," *Journal of Applied Science and Technology Trends*, vol. 2, no. 1, pp. 10–19, 2021. <https://doi.org/10.38094/jastt20179>
- [10] R. Chaganti, 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*, vol. 14, no. 16, p. 3914, 2022. <https://doi.org/10.3390/cancers14163914>
- [11] A. I. Dawood, Q. Q. Thabit, and T. O. Fahad, "Thyroid disease prediction with machine learning algorithms," *Eurasian Research Bulletin*, vol. 18, pp. 229–237, 2023.
- [12] L. Aversano *et al.*, "Thyroid disease treatment prediction with machine learning approaches," *Procedia Computer Science*, vol. 192, pp. 1031–1040, 2021. <https://doi.org/10.1016/j.procs.2021.08.106>
- [13] K. Shankar, S. K. Lakshmanaprabu, D. Gupta, A. Maselena, and V. H. C. De Albuquerque, "Optimal feature-based multi-kernel SVM approach for thyroid disease classification," *The Journal of Supercomputing*, vol. 76, pp. 1128–1143, 2020. <https://doi.org/10.1007/s11227-018-2469-4>
- [14] J. Gu *et al.*, "A machine learning-based approach to predicting the malignant and metastasis of thyroid cancer," *Frontiers in Oncology*, vol. 12, 2022. <https://doi.org/10.3389/fonc.2022.938292>
- [15] D. Bhende *et al.*, "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)*, 2024, pp. 1–6. <https://doi.org/10.1109/SCEECS61402.2024.10481980>
- [16] M. Faiz and S. W. A. Rizvi, "Thyroid disease detection using machine learning," *Journal of Applied Science and Education (JASE)*, vol. 3, no. 2, pp. 1–9, 2023. <https://doi.org/10.54060/jase.v3i2.32>
- [17] A. K. Chaudhuri, D. K. Banerjee, A. Das, and A. Ray, "A multi-stage approach combining feature selection with machine learning techniques for higher prediction reliability and accuracy in heart disease diagnosis," *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 7, 2021.
- [18] M. Riajuliislam, K. Z. Rahim, and A. 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)*, 2021, pp. 60–64. <https://doi.org/10.1109/ICICT4SD50815.2021.9397052>

- [19] A. Sultana and R. Islam, "Machine learning framework with feature selection approaches for thyroid disease classification and associated risk factors identification," *Journal of Electrical Systems and Information Technology*, vol. 10, 2023. <https://doi.org/10.1186/s43067-023-00101-5>

7 AUTHORS

Deepali Bhende has earned a Master's degree in Computer Science from G. H. Rasoni Institute of Information Technology, Nagpur. Presently, she is actively pursuing her Ph.D. at G. H. Rasoni University, Saikheda (MP). She has made notable contributions through several published papers in various international journals, with seven papers indexed in SCOPUS. Furthermore, she has successfully obtained a patent and a copyright, demonstrating her innovative contributions to the field (E-mail: deepali.bhende.phdcs@ghru.edu.in).

Dr. Gopal Sakarkar is an experienced academician, holds a master's degree (2006) and a Ph.D. (2017) from S.G. B. Amravati University, Amravati, with over 15 years of teaching and research expertise. As a research guide, he mentors Ph.D. scholars, and his prolific research output comprises 55+ papers in international journals and conferences, including SCI, ESCI, WoS, Springer, Elsevier, IEEE, AIP, and SCOPUS indexed conferences. With 200+ Google Scholar citations, he holds a 6 h-index, a 5 i10 index, and 79 citations in Scopus with a 3 h-index.

Ambika Jaiswal is an Assistant Professor of Computer Science at SGB Amravati University with 14 years of postgraduate teaching experience. She is pursuing a Ph.D. in Computer Science, focusing on mobile cloud computing and AI/ML. She has published extensively in high-impact journals and conferences, and her research has led to patents in India and in UK. She is committed to advancing her field through innovative research in mobile cloud computing and AI/ML.

Punam Khandar has completed her M. Tech (CSE) degree from RTMNU, Nagpur, in the year 2015 and currently serves as an Assistant Professor at Shri Ramdeobaba College of Engineering and Management, Nagpur. She is presently pursuing a Ph.D. in Computer Science and Engineering at KIIT, deemed to be University, Bhubaneswar. Her research interests encompass Machine Learning, Deep Learning, and Computer Vision. She has recently authored papers on various topics, including a survey on Decision Trees, detection of leaf diseases, identification of heart diseases, and the use of X-rays for identifying COVID-19.

Satyajit Uparkar, a certified data scientist, has been working as an Assistant Professor at Shri Ramdeobaba College of Engineering and Management, Nagpur, for the past 11 years. Holding three postgraduate degrees, he specializes in data analytics with a focus on Data Mining, Scalable Data Science, and Operation Research Modeling. He has received two best paper awards at international conferences and has contributed 30 research papers to various national and international journals. Alongside his academic role, he also serves as a data science consultant for local companies.

Dr. Lalit Agrawal is an Assistant Professor at the School of Computer Engineering, Ramdeobaba University, Nagpur. He holds a Ph.D. in Computer Science and Engineering from RTMNU, Nagpur, with 11 years of teaching and research experience. He has published extensively in reputed journals indexed in SCI, Scopus, and Web of Science, and has won a best paper award at an international conference. His research interests include data mining and algorithm optimization.