Vol. 3 No. 2 (2025)

Spectrum of Engineering Scien

SPECTRUM OF ENGINEERING SCIENCES

Online ISSN
3007-3138
Print ISSN
3007-312X

Robust Analysis of Hypothyroidism Detection Using Ensemble Modeling Techniques

Sohaib Latif[1*]
Department of Computer Science and Software Engineering, Grand Asian University, Sialkot, 51310, Pakistan. sohaiblatif095@gmail.com
Sadia Karim[2]
Department of Computer Science, The University of Chenab, Gujrat, 50700, Pakistan.sadiamobeen92@gmail.com
Daniyal Affandi[3]
Department of Computer Science, The University of Chenab, Gujrat, 50700.
affandidaniyal305@gmail.com

Abstract

Hypothyroidism is a prevalent medical condition that demands timely detection and cure because early detection can contribute to decreasing death rates. To check the thyroid level in the human body current diagnostic methods, for example, TSH, T3, FTI, etc are performed, which places a heavy workload on paramedical and healthcare professionals. This article dives deep into the application of Machine Learning algorithms for the detection and classification of hypothyroidism automatically. To develop such a system it is essential to increase the precision rate and decrease the error rates of the ML model. This study was done on six ML classifiers J48, Random Forest, Random Tree, Naïve Bays Multiclass and MLP, which produced significant results for their comparative analysis. After evaluation, it was revealed that models built up to predict the issue using a MLP and random Tree generated the highest level of accuracy, 99.9565% and 99.8692% respectively. On the other hand, it was also seen that the time MLP took was way more than any other model. So far, our research shows the best results for MLP Accuracy and the lowest error rate. This effort provides vital insights into the possible Machine learning

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

techniques for the detection and classification of Hypothyroidism making a pathway for more accurate and realistic diagnostic tools in clinical practices.

Keywords: Hypothyroidism, Ensemble Techniques, MLP, Naive Bayes, Random Forest, Multiclass, J48

Introduction

Thyroid is a little gland that is in the shape of butterfly observed in the upper front aspect of neck. It's responsible for generating certain hormones that control the body's different functions and energy consumption. Those hormones have tremendous effects on the body's every organ and most critical functions of body. For example, to regulate respiration, managing weight , digestion and so on. If any issue arises within this gland it causes thyroid diseases, which come under the endocrine system. TSH is a thyroid - stimulating hormone that pinpoints whether this thyroid gland works properly or not. It represents that how this thyroid gland functions: under-active or overactive. If it is high it is termed as Hyperthyroid and if it is produced lesser than needed it is called Hypothyroid. It reveals how abundant this hormone is present in blood, produced by a body's pituitary gland in the brain, telling the thyroid gland to release more hormones in accordance with the needs of the body.

Hypothyroidism affects 4.1% of the population in Pakistan, whereas hyperthyroidism affects 5.1%. Approximately, 2% of United Kingdom population is suffering from this critical clinical illness and 4.6% in the United States. Iodine deficiency plays a major role in spreading this problem. There are different factors additionally which can lead to hypothyroidism which is the substantial use of pesticides, endocrine disorder, and contaminated and unhygienic drinking water. These issues are common place, easily identifiable, and normally manageable, but, if left untreated, extreme consequences may result, ranging from the relatively simple matter of an enlarged goiter to potentially life-threatening conditions such as cancers. Hyperthyroidism is caused by overproduction of hormone, even as hypothyroidism is a end result of underproduction. In

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

spite of the distress they may cause, accurate analysis and treatment are available. Globally, thyroid problems vary from gender, age, race, and area, mostly depending on the changes in the intake of iodine. Approximately one-third of the world's populace lives in places that have an iodine deficiency, a main cause of global thyroid disorder.

Timely analysis of hypothyroidism plays a vital role in preventing various problems like joint pain, coronary heart issues, and obesity. This problem is definitely curable and, in a few instances, general reversal from hypothyroid is visible. In women, hypothyroidism can cause infertility (not able to conceive a baby), leading to the lifestyles-threatening situation called myxedema. Early detection, which is supported in this study, is highly important for saving lives, reducing mortality ratio, inspiring and guiding clinicians using available datasets. This paper advises if a patient has hypothyroidism, hyperthyroidism, or euthyroid {having standard levels of TSH hormone}. It describes a machine learning-based methodology to capture measurements of Thyroid hormone and other relevant information with respect to the patient and generate the output. At present Machine learning is playing a pivotal role in the field of data science analytics to cope with the variety of mental and physical illness. It proves quite efficient in predicting thyroid conditions, detecting pneumonia, and forecasting consequences related to COVID-19.

Fig.1 shows the latest data of hypothyroidism 2022-2023 occurrence of hypothyroidism across different countries. The average number of Hypothyroidism patients reported in USA is 4.6% while in India, Pakistan, China, and Europe and have reported 10% ,4.1%,13.95% and 5% respectively.

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

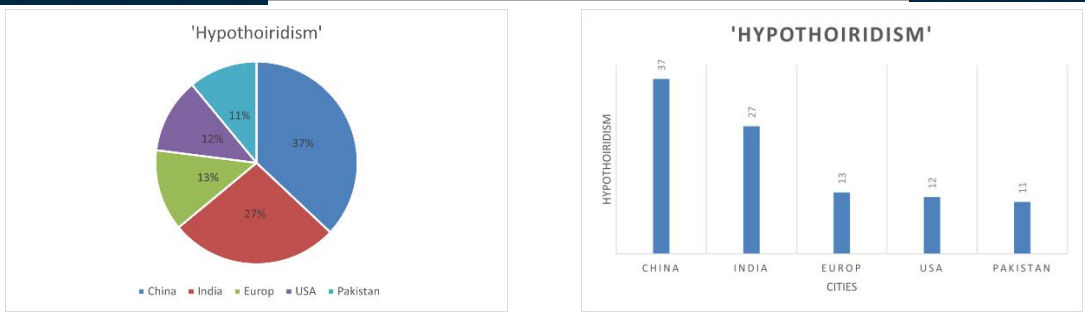

Figure 1: Incidence Of Hypothyroidism In Different Countries

Literature Review

In the past many researchers have already worked on different medical data sets and verified the accuracy of machine leaning algorithms (Shiuh, Tong Lim, et al. 2023). They used features wiz library in python to finalize the best models among various ML models and algorithms. They experimented on the data set with and without feature selection. When used with performing feature selection Random Forest was seen best while Boost showed best results without feature selection (Alyas, Tahir, et al. 2022). Study was done to improve and aid in healthcare system by evaluating the sampled as well as un-sampled data set. A manipulative study shows that highest accuracy was achieved by Random Forest with 94.8% (Alnaggar, Mona, et al. 2023). Results of various algorithms in Thyroid detection show that if XGBoost is used after optimization of parameters it gives the highest level of accuracy and precision.

(Mir, Yasir Iqbal, and Sonu Mittal. 2020) used a different approach for thyroid detection called hybrid model in which they used a real-time data set. Bagging was considered the best approach considering all the features of the dataset. SVM was found best if only pathological features of the data set were taken into account and finally J48 showed remarkable results on serological features of the data set (Latif, Sohaib, et al. 2023). Uses quite different approach for detecting thyroid problem the body. They used the data set and applied various ML algorithms on the whole data set and later they removed three attributes (query thyroxin, query hypothyroid, query hyperthyroid) of the given data and then compared the accuracy of both

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

the approaches. it was found that results were better in shape if three attributes were removed. Overall SVM showed the best accuracy.

(Naeem, Awad Bin, et al. 2023) experimented on KNN, Naive Bays and SVM to predict the level of hypothyroidism by using Rapid Miner tool with four interfaces of the tool were used and they found SVM was the number one approach with 84.72% accuracy. To reduce the wrong or delayed diagnosis of thyroid. (Kalyani, B. J. D., et al. 2024) used CNN t detect the exact type of thyroid irregulars the body. This model works on the CT scans and ultrasound Images, showing results of 97.2 and 94.2 of accuracy resp. According to (Aversano, Lerina, et al. 2023) old and outdated data sets are being used for detecting thyroid problem but they are producing results frequently and efficiently. Mostly Neural networks are most commonly used AI technique in this regard. In a study (Abbad Ur Rehman, Hafiz, et al. 2021) they applied a number of ML algorithms to detect thyroid disease. Some were applied with feature selection, and some were not. Important thing is that they used their own local hosp. data set and added three additional attributes as well to enhance the accuracy of the results. L1 and L2 feature selection were used and L1 showed significantly correct results as compared to L2 or no feature selection approach.

According to (Sha, Mohemmed. 2023) misclassified results can cause low or zero accuracy in predicting the thyroid disease. They proposed optimized features selection method of QPSO instead of traditional PSO, also QSVM instead of Traditionally used SVM methods. Paper proposes that this combination provides the desired accuracy. The study uses the supervised and unsupervised ML models and trained the models to minimize chances of FP and FN. Random Forest and standard scalar methods were applied. (Mousavi, Seyedeh Somayeh Zarin, Morteza Mohammadi Zanjireh, and Marzieh Oghbaie. 2020) children with thyroid issue can have a outclass treatment if diagnose and screened without error. They discuss in their paper that if SVM along with bagging technique is apply in combination it can cause great results with accuracy also they

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

discussed other factors like family history of illness, family marriage, type of delivery etc. also plays a major role in thyroid prevalence in infants.

(Saleh, Dhekre Saber, and MohdShahizan Othman. 2023) variable symptoms of Thyroid disease make it difficult to diagnose that why treatment often delayed. this review explain that imbalance and unjust detection causes hindrances in achieving accuracy. According to this study overall neural network and RF models are used in most of the papers. (Chaganti, Rajasekhar, et al. 2023) claims that mostly used thyroid detection methods and ML algorithm are focusing on optimizing the model instead of features manipulation. This study presents the missing feature part. They performed various feature optimizations to enhance the results. Hashimoto's thyroid can be predicted using this model. They used extra tree classification model and attained highest accuracy with RF model. (Islam, Saima Sharleen, et al. 2022) compared eleven algorithm from statistical, neural and tree algorithms from Machine learning. considering the main goal detecting thyroid presence, ANN was found best in term of accuracy(Hosseinzadeh, Mehdi, et al. 20221). Recently internet of medical things has made a significant position in the world of medical informatics. ANN and Internet of medical things (IoMT) were used in combination with a new proposed MMLP with the ability to back propagate. Authors showed the superiority of the model by using a set of six networks of MMLP (Sultana, Azrin, and Rakibul Islam. 2023). This study employs several Ml algorithms for TD (thyroid detection) by apply different feature selection methods. Finally, they concluded that RF if combined with LASSO technique was considered the best with 99 % of accuracy.

(Savcı, Ege, and Fidan Nuriyeva. 2022)presents a web app creation project in which user enter data about the symptoms of thyroids, later various ML algorithms and regression are applied to train the data which work at beck end and talk about the diagnosis of the thyroid illness. (Frye, Maik, Johannes Mohren, and Robert H. Schmitt. 2021)found the best results

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

were obtained when correlation extraction was applied on various ML algorithms. Among which the best one was ANN, SVM and KNN.

Material And Methods

In this study, a predictive model was developed for the diagnosis of thyroid disorders by applying supervised learning techniques. The hypothyroidism dataset is made up of 30 attributes and has 3,772 instances, derived from the data repository of the WEKA simulator, which is widely used for machine learning experiments and research. WEKA is an entire collection of machine learning algorithms and tools for data mining activities. It provides various functionalities for data preprocessing, classification, regression, clustering, association rule mining, etc. The experiments in the current study were performed on WEKA version 3.8.6 that provides a user-friendly interface for performing complex data analysis.

In this research work, supervised learning algorithms include Naive Bayes, Multilayer Perceptron (MLP), Random Forest, Random Tree, J48-a decision tree algorithm-and Multiclass Classification. These mentioned algorithms are tested using the hypothyroidism dataset to identify hypothyroidism in terms of clinical attributes. This is achieved through accuracy, efficiency, and the generalization of algorithms toward unseen data to compare various methods of diagnosis related to thyroid disorders.

Hypothyroidism Prediction Model Development

ML gives a computer or machine the capability to learn and improve on the performance without direct programming. The process of creating an ML model entails applying various classification techniques in determining meaningful outcomes from data. In the contemporary world, the sector that ML can apply includes robotics and health diagnosis as it provides ways to eliminate repeated tests while maximizing accuracy in diagnosis. In the healthcare sector, for instance, the ML models are very effective in disease diagnosis, prognosis of patients' conditions, and tailoring a suitable treatment plan for each patient.

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Figure 2 illustrates the modeling of the suggested predictive model with several steps including collection of data, preprocessing step, feature selection, training of model, and testing. The former process involves data collection from proper sources, which ensures that it portrays the problem domain. Under the preprocessing step, raw data is cleaned and transformed in order to handle missing values, normalization, data transformation, and so on.

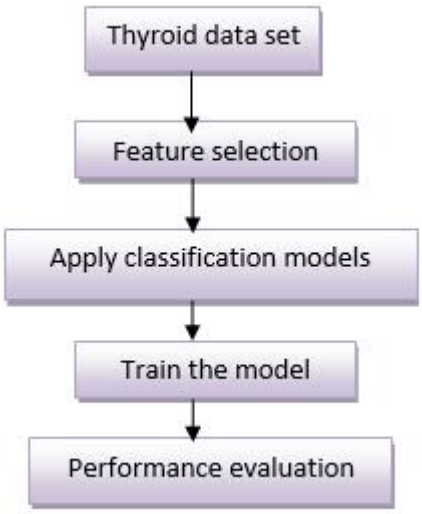

Figure 2: Methodology

Use of Machine Learning is so much rapid in today's world due to its greater impact on problem solving, but along with that there is a dire need of cleaning the poor quality of available data (Pal, Madhumita, Smita Parija, and Ganapati Panda. 2022). Data preprocessing (Kang, Myeongsu, and Jing Tian. 2018) is the first process done in Machine Learning (ML) techniques. It is consisting of various activities like, cleaning the dirty or noisy data, normalizing the unnormalized data, filling the missing values or putting mean value of attribute, transforming raw data into an acceptable and useful form. Corrupt or noisy data, containing outliers or missing values may interfere with the output ML model. Binning methods are utilized to remove such disturbances from the dataset. Feature selection identifies the

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

most relevant attributes or features with an aim of improving the efficiency and performance of the model and reducing overfitting cases. In this stage, the model is trained on a subset of data, where it learns from patterns and relationships in data. While training, adjustments to internal parameters of an algorithm are made in efforts of reducing errors of prediction. The final stage is testing, wherein evaluation on unseen data tests for how well the model is performing in terms of generality, accuracy, and capability to predict things more closely. The cycle is repetitive wherein every round fine-tunes a particular model to yield an ideal solution. The cycle repeats till the best performance can be obtained from one or more of these models developed for predictive modeling. All these are key stages of developing a predictive model that would guarantee reliability of ML models in a practical application.

After preprocessing different filters were applied on the data set like, smote (Mukherjee, Mimi, and MatloobKhushi. 2021), remove percentages and later feature selection method was applied on clean data to see the key attributes which are needed to draw the accurate and best results for our thyroid data set. In the study presented, this step is undertaken to eliminate redundant data, hence, improving accuracy and other important performance parameters, named as recall, and precision while also minimizing time for computation . Feature selection is accomplished using various available methods, including embedded methods, filter methods, and wrapper methods. Lastly, the classification results have been found using accuracy, precision, recall, F1-score, ROC .

The multilayer perceptron (MLP), a Artificial Neural Network called feed-forward, that means that if there is change in the output of the neuron there is no change on the neuron itself. It mimics same like human brain's organization patterns. It consists of many layers hidden between the input and output layers, with the number of hidden layers determined by the specific data mining task. Connected Neurons in the hidden layers makes connection and these are called weights, are continually updated during a

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

learning phase. The learning phase iterates until the error falls below a set threshold.

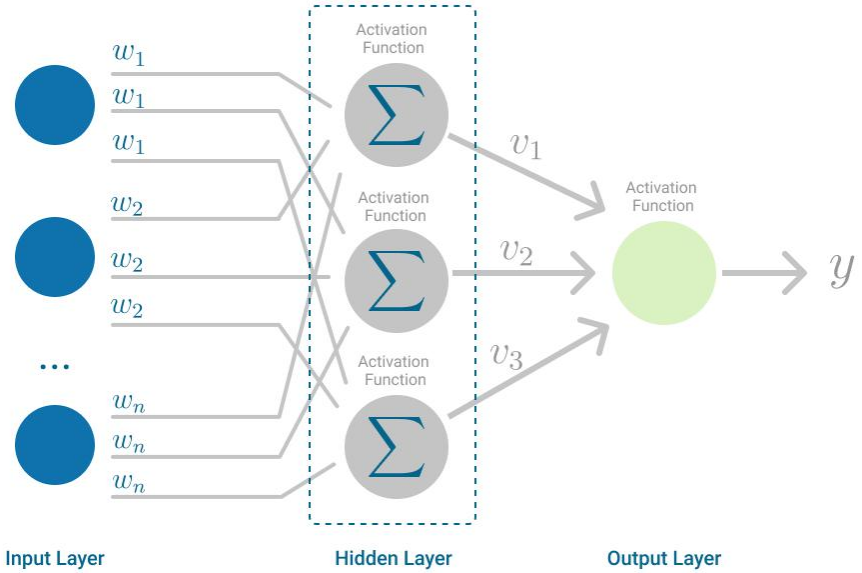

Figure 3: Artificial Neural Network Model Formation

One of the most robust machine learning algorithms developed by Leo Breiman and Adele Cutler is called Random Forest. In this approach, the phenomenon of "bagging" is followed to reduce overfitting and hence increase the accuracy of predictions compared to individual decision tree. It can easily determine the most important variables in a dataset.Multiclass Classification is a specific class of machine learning model that helps classify data on the basis of more than two classes or categories. These can classify a single data point to more than one unique class. For instance, in healthcare, multiclass classification models are used in the diagnosis of several diseases from a set of symptoms or test results.

Naïve Bayes is a probabilistic classifier that uses Bayes' Theorem for forcasting the probability of a class given a set of features. The word "naïve" comes from its assumption of conditional independence among the features or variables, which makes the simplification not strictly correct in most real cases. This model also has the advantage of computation. It can

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

be very quick to train and to use, making it well suited to real-time applications, when a large amount of data is to be processed speedily. its performance using limited computational resources and resistance to noisy data make it a good choice for most applications involving practical applications in which assumptions of independence are approximately held.

Results And Discussion

WEKA is a JAVA language tool for making use of variety of supervised or unsupervised machine learning (ML) algorithms. The platform facilitates visualization environment and allows for the incorporation of new ML algorithms. In WEKA 10-fold cross-validation can be applied. It is used for different Data Mining techniques such as Classification, Association Data Preprocessing, Clustering, and also Visualization. This paper is an attempt to acquaint the researcher with the WEKA interface Performance evaluation is conducted using key metrics such as recall, accuracy, ROC, F1-score

Result Prediction

This section offers the insight into the proposed model to detect the level and class of hypothyroidism results. Fig 4-9 shows the performance results of random forest, j48, naive bays, multiclass classifier, random tree and multi-level perceptron (MLP) respectively telling about the level of precision, , recall value , F1-score, level of accuracy and ROC value for each class, i.e., t (tsh measure true) and f (tsh measure false) in the dataset. The results then demonstrate that for every t and f class, the MLP model, shows the highest among all precision that is 1, which is above all high when compared to other suggested models. Similarly, when taking into account the value of recall, F1-score, and ROC, results of MLP model will be considered better than other models which were chosen for present study . The effectiveness of any model is dependent upon demonstrating a high level of accuracy within a brief timeframe. In this study, time is considered a critical and trust worthy perameters for judging the performance of the proposed model. It serves as a crucial parameter for identifying the most suitable model for a given problem.

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

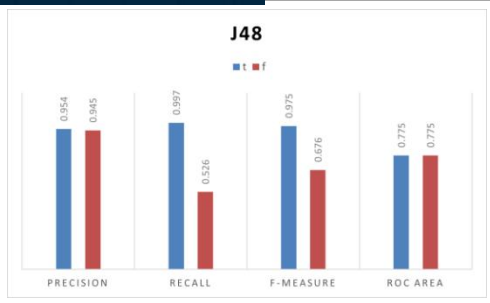

Figure 4: J48 Result

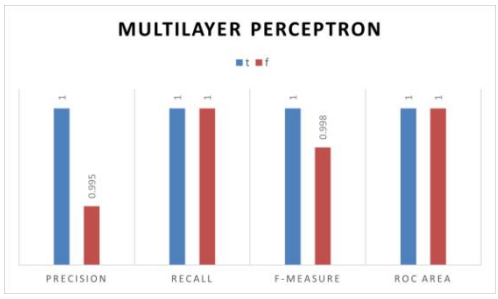

Figure 5: Multilayer Perceptron

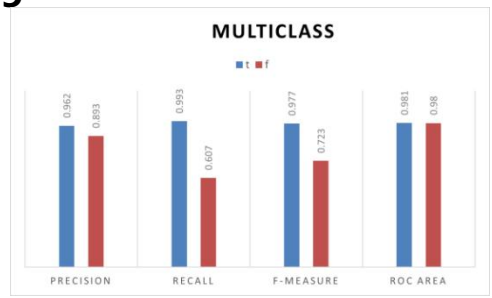

Figure 6: Multi-Class

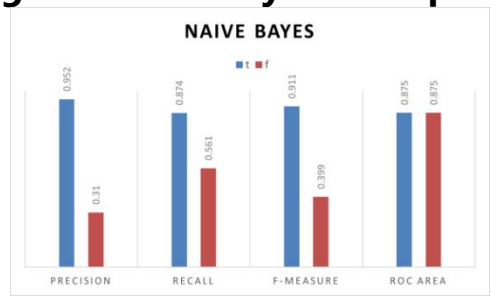

Figure 7: Naïve Bayes

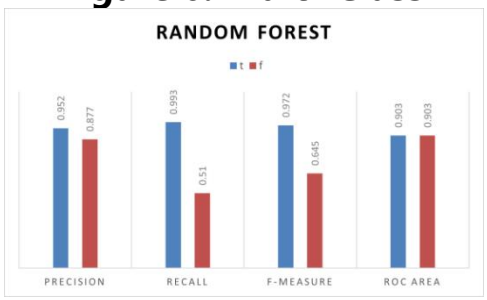

Figure 8: Random Forest

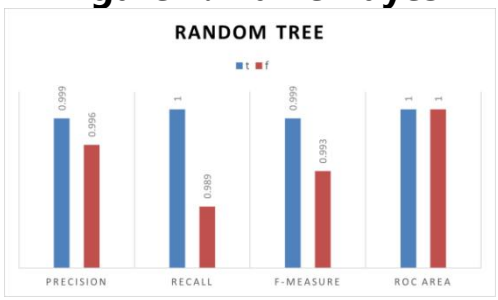

Figure 9: Random Tree

While taking Recall into account, F- measure and precision into account Following Fig 10, 12 and 12 demonstrates that naïve bays get lowest score among all models, while MLP and Random Tree shows no fluctuation for both classes (t, f).

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

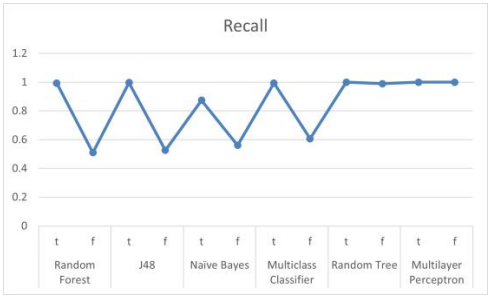

Figure 10: Re-Call Comparison

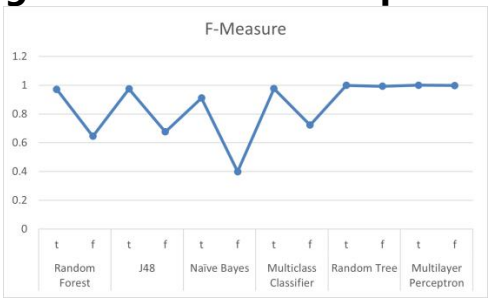

Figure 11: F-Measure Comparison

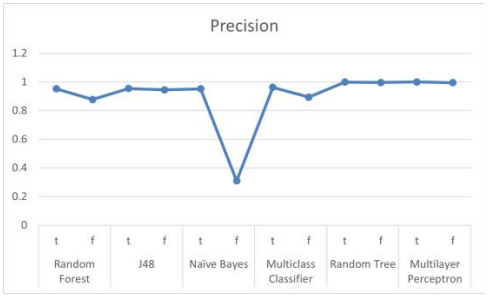

Figure 12: Precision Comparison

Time Assessment For Each Algorithm

Consequently, the results are also assessed in terms of the time required to execute the complete model, as depicted in Figure 13. The comparison reveals that the random tree classifier exhibits the shortest runtime in the experiment, while the multilevel perceptron model is identified as the most time-consuming. The order of time consumption for predicting hypothyroidism is found to be multilayer perceptron, random forest, multiclass classifier, j48, random tree and naïve Bayes ranging from the better to moderate time-consuming models.

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

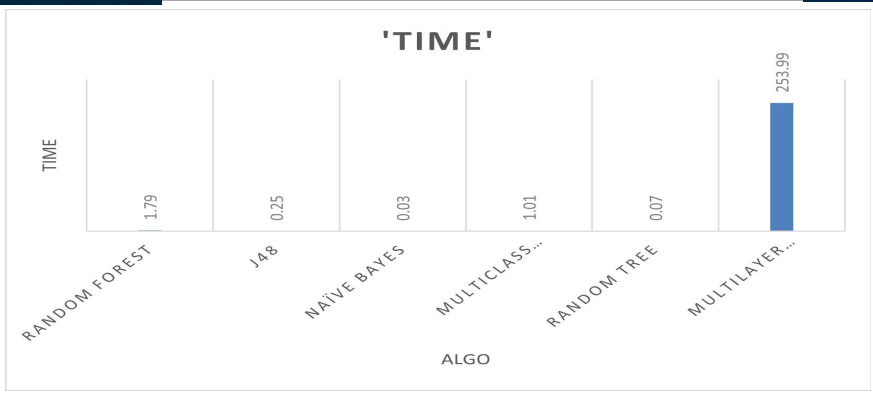

Figure 13: Compilation Time Taken By Various Ml Models

Total Comparison Of Algorithm

Table 1 shows the working performance of a mixture of prediction models for thyroid. Two divisions for hypothyroidism were made in this model first was t division and second was f division. The classifier named naïve bays provided an average accuracy of 84.4854% whereas general error rate of 0.3541. The exhibition of error for Random Forest was 0.208 with a rate of accuracy value of 94.843% which is slightly higher than the naive Bayes classifier. When we consider error value of the multiclass model it is showing value of 0.1695, it is depicting an accuracy of 95.7261%.

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

Table 1: Total Comparison

Classifier	Naive Bayes	J48	Multiclass Classifier	Random Tree	Random Forest	Multi-LayerPerceptron MLP
Accuracy (%)	84.48	95.36	95.72	99.86	94.84	99.95
Error Rate	0.3541	0.2081	0.1695	0.0347	0.208	0.019
F1 Score	0.864	0.948	0.954	0.999	0.942	1
Roc Curve	0.875	0.775	0.981	1.000	0.903	1
Recall	0.845	0.954	0.957	0.999	0.948	1
Precision	0.893	0.953	0.955	0.999	0.946	1

Error And Accuracy Analysis

Random tree model's error rate is identified as 0.0347, while its accuracy level is 99.86%. This model shows good results and it is better than the all others except multilayer perceptron. J48 has the accuracy of 95.36% while 0.2081 is error. Lastly, the MLP Classifier shows the error rate as (0.019) with an accuracy value of 99.95% which is a competitively best score [fig 14]. As a result, two models, namely random tree, and MLP, have been identified as the best performing models with the highest accuracy values as 99.86%and 99.95% respectively, as shown in fig 15.

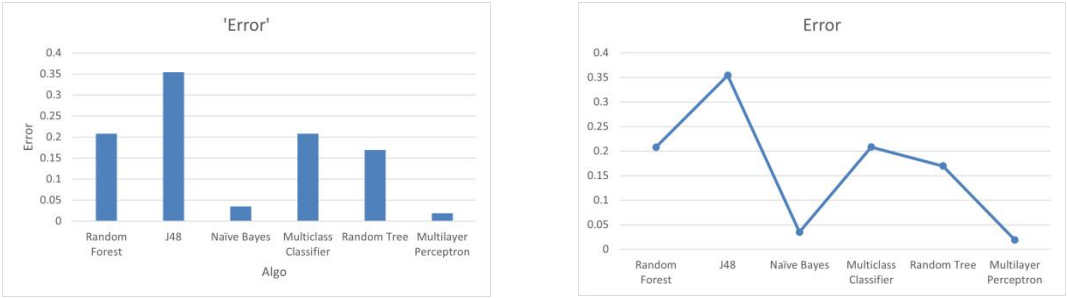

Figure 14: Comparison Of Error In All Algorithms

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

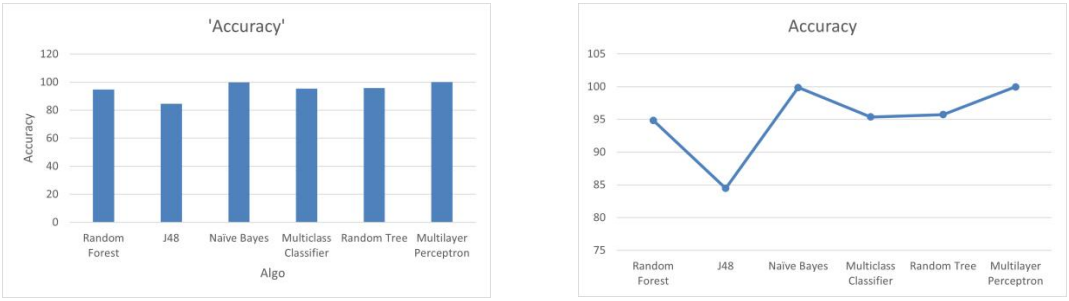

Figure 15: Compilation Of Accuracy

Overall Results Of Prediction Comparison

Finally there is a comparison to predict about the disease . we took into account the parameters of value of recall, F1 value, considered ROC and precision use Weka tool, the result was that multilayered perceptron is performing best. Random tree is better than others in case of, recall, F1-score, precision and ROC while multiclass is better than j48. naïve bays performance is considerably low in this regard as shown in Fig 15 and Fig 16. It is safe to say that , by taking into account the valule of precision, F1-score, recall and ROC, the MLP (Multilayered percepteron) is better . Taking time into consideration, the Random tree model showing trustworthy results . Because MLP takes a much longer time but If we ignore time MLP outperforms all algorithms to forecast thyroid detection.

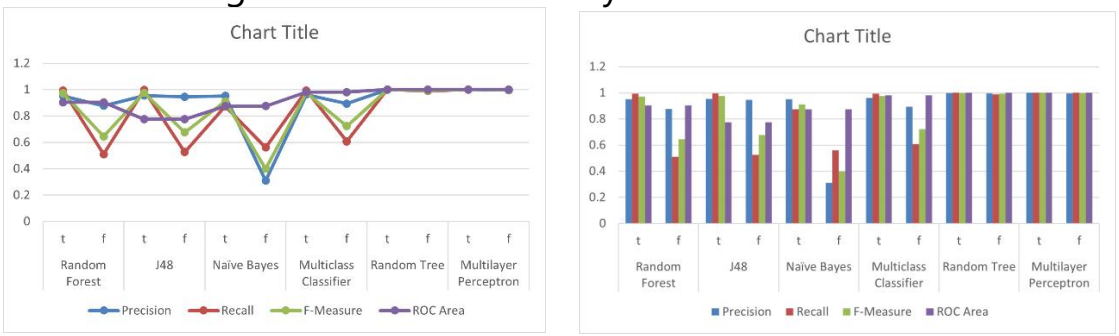

Figure 16: Total Comparison Of All Algorithms With Respect To (t, f)

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

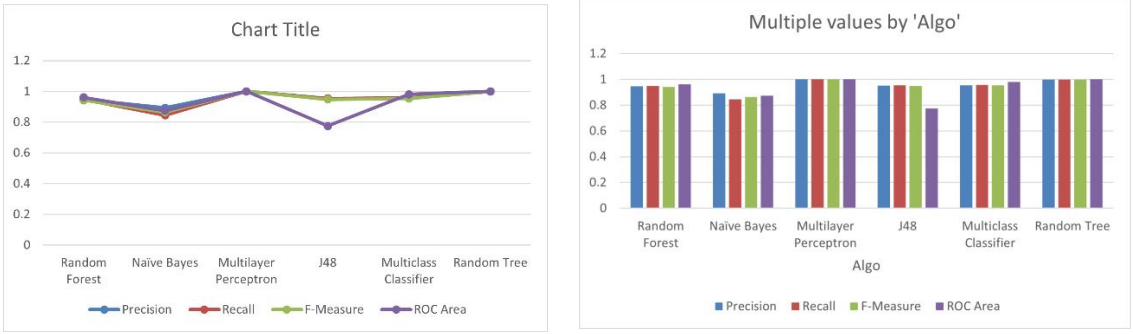

Figure 17: Total Comparison Of All Algorithms

Conclusion And Future Work

Hypothyroidism, a condition caused by an underactive or malfunctioning thyroid gland, is spreading increasingly all over the world, and more women than men are its victims. In this study, a prediction approach using machine learning techniques was developed to identify hypothyroidism by employing various algorithms, including Random Forest (RF), Multiclass Classifier, Naïve Bayes, Random Tree, J48, and Multilayer Perceptron (MLP). Each of the algorithms was tested stringently for their correctness in predicting hypothyroidism; the performance metrics accuracy, precision, recall, F1-score, and ROC values are calculated to determine the merits of each of the models designed. Through careful review, it can be observed that the MLP and Random Tree-based models reached an accuracy level of 99.95% and 99.86% respectively. Of the above, the MLP was the best model. However, it is also relatively long to execute because the model demands extensive training data to achieve high accuracy, which tends to make it more computationally intensive than other algorithms tested. This means that while MLP is a stronger predictor, it may not be the best choice for situations where both computing power and time are scarce.

As such, this study includes a comparative analysis study with existing hypothyroidism prediction models, where the proposed models, specifically MLP, outperform previous methods, suggesting that accuracy and reliability are improved for diagnosing hypothyroidism. Such findings are extremely useful for healthcare applications. The model developed by this research

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

study will further be generalized and applied for other prediction needs beyond hypothyroidism prediction. Suitable adaptation of such ML models and the right set of medical datasets will promise them for supporting diagnosis and classification in other diseases too. The tools will, therefore, benefit doctors and clinicians with valid and data-driven support for decisions leading to an improved process for decision-making in medical diagnostics as well as to better disease-identification accuracy across domains. With ML models now quite proactively contributing to the prevention and management of diseases at earlier phases, this step might form one of the leading breakthroughs toward predictive health care.

References

1. Latif, S., 2024. Robust Decision Support System for Stress Prediction Using Ensemble Techniques. Journal of Innovative Computing and Emerging Technologies, 4(2).
2. Alyas, Tahir, Muhammad Hamid, Khalid Alissa, Tauqeer Faiz, Nadia Tabassum, and Aqeel Ahmad. "Empirical method for thyroid disease classification using a machine learning approach." BioMed Research International (2022).
3. Alnaggar, Mona, Mohamed Handosa, Tamer Medhat, and M. Z Rashad. "Thyroid disease multi-class classification based on optimized gradient boosting model." Egyptian Journal of Artificial Intelligence 2, no. 1 (2023): 1-14.
4. Mir, Yasir Iqbal, and Sonu Mittal. "Thyroid disease prediction using hybrid machine learning techniques: An effective framework." International Journal of Scientific & Technology Research 9, no. 2 (2020): 2868-2874.
5. Latif, Sohaib, Xian Wen Fang, KaleemArshid, Abdullah Almuhaimeed, Azhar Imran, and Mansoor Alghamdi. "Analysis of birth data using ensemble modeling techniques." Applied Artificial Intelligence 37, no. 1 (2023): 2158273.
6. Latif, S., Fang, X., Mohsin, S.M., Akber, S.M.A., Aslam, S., Mujlid, H. and Ullah, K., 2023. An enhanced virtual cord protocol based multi-casting

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

strategy for the effective and efficient management of mobile ad hoc networks. Computers, 12(1), p.21.

7. Kalyani, B. J. D., KopparthiBhanuPrasanth, Pannala Krishna Murthy, S. Kavitha, and S. Neelima. "Revolutionizing Thyroid Disease Forecasting with API Enhanced Convolutional Neural Networks." International Journal of Intelligent Systems and Applications in Engineering 12, no. 6s (2024): 64-68.

8. Latif, S., Fang, X.W., Arshid, K., Almuhaimeed, A., Imran, A. and Alghamdi, M., 2023. Analysis of birth data using ensemble modeling techniques. Applied Artificial Intelligence, 37(1), p.2158273.

9. Abbad Ur Rehman, Hafiz, Chyi-Yeu Lin, ZohaibMushtaq, and Shun-Feng Su. "Performance analysis of machine learning algorithms for thyroid disease." Arabian Journal for Science and Engineering (2021): 1-13.

10. Sha, Mohemmed. "Quantum intelligence in medicine: Empowering thyroid disease prediction through advanced machine learning." IET Quantum Communication (2023).

11. Viswanatha, V. "Thyroid Disease Detection Using Machine Learning approach." (2023).

12. Mousavi, SeyedehSomayehZarin, MortezaMohammadiZanjireh, and MarziehOghbaie. "Applying computational classification methods to diagnose Congenital Hypothyroidism: A comparative study." Informatics in Medicine Unlocked 18 (2020): 100281.

13. Saleh, Dhekre Saber, and MohdShahizan Othman. "Exploring the Challenges of Diagnosing Thyroid Disease with Imbalanced Data and Machine Learning: A Systematic Literature Review." Baghdad Science Journal (2023).

14. 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.

15. Islam, Saima Sharleen, Md Samiul Haque, M. SaefUllahMiah, Talha Bin Sarwar, and RamdhanNugraha. "Application of machine learning algorithms

Spectrum of Engineering Scien
SPECTRUM OF ENGINEERING SCIENCES
Online ISSN
3007-3138
Print ISSN
3007-312X

to predict the thyroid disease risk: an experimental comparative study." PeerJ Computer Science 8 (2022): e898.

16. Latif, S., 2024. Robust Decision Support System for Stress Prediction Using Ensemble Techniques. Journal of Innovative Computing and Emerging Technologies, 4(2).

17. 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.Vasan, Chandan R. Chetan, Bangalore DSU, Chethan MS, and H. S. Devikarani. "Thyroid detection using machine learning." Computing (PDGC-2018) 20 (2018): 22.

18. Savcı, Ege, and Fidan Nuriyeva. "Diagnosis of Thyroid Disease Using Machine Learning Techniques." Journal of Modern Technology & Engineering 7, no. 2 (2022).

19. Latif, S. and Mussarrat, N., 2024. Impact of Digital Technologies on Consumer E-Purchasing Decisions Using Ensemble Techniques. International Journal of Information Systems and Computer Technologies, 4(1), pp.10-19.

20. Pal, Madhumita, Smita Parija, and Ganapati Panda. "Enhanced Prediction of Thyroid Disease Using Machine Learning Method." In 2022 IEEE VLSI Device Circuit and System (VLSI DCS), pp. 199-204. IEEE, 2022.

21. Kang, Myeongsu, and Jing Tian. "Machine Learning: Data Pre-processing." In Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things, 111-130. 2018.

22. Mukherjee, Mimi, and MatloobKhushi. "SMOTE-ENC: A Novel SMOTE-Based Method to Generate Synthetic Data for Nominal and Continuous Features." Applied System Innovation 4, no. 1 (2021): 18.