How to Cite:
Singh, T., Sahu, A. K., Dubey, S., Sharma, M. P., Verma, S., & Kumar, C. (2022). Treatment of thyroid disease through machine learning predictive model. *International Journal of Health Sciences*, *6*(S8), 3176–3188.
https://doi.org/10.53730/ijhs.v6nS8.12813

Treatment of thyroid disease through machine learning predictive model

Tanvir Singh
IIMT College of Engineering, Greater Noida India, Assistant Professor, Department of MCA
Corresponding author email: ulab12345@gmail.com

Dr. Ajay Kumar Sahu
Greater Noida Institute of Technology (GNIOT), Greater Noida India, Associate Professor, Department of IT
Email: ajay4989@gmail.com

Dr. Shivani Dubey
Greater Noida Institute of Technology (GNIOT), Greater Noida India, Associate Professor, Department of AI-ML & IOT
Email: dubey.shivani@gmail.com

Dr. Mahendra Prasad Sharma
IIMT College of Engineering, Greater Noida India, Associate Professor, Department of IT
Email: mahendrasharma.gn@gmail.com

Shanu Verma
Pranveer Singh Institute of Technology, Kanpur India, Assistant Professor, Department of CSE
Email: shanu.verma56@gmail.com

Chaman Kumar
IIMT College Of Engineering, Greater Noida India, Assistant Professor, Department of IT
Email: chamankumar31@gmail.com

***Abstract*---**The thyroid seems to be an part of the endocrine system that is placed toward the front of neck and produces thyroxine, which are essential for our overall health. If it fails, thyroid hormone production will either be insufficient or excessive. Machine learning techniques and data mining are critical in processing large amounts of data, particularly in the health care system, where there has been a massive amount of information and data need to be managed. In our

research on thyroid disease, we used machine learning approaches. In our study, we used statistics from patients, a few of which has hyperactive thyroid glands moreover those have hypothyroidisms; therefore, overall algorithms were used. These study aims to divide this disease in few categories like as hypothyroidism, regular and hyperthyroidism. Support vector machine include KNN, naive-bayes, logistic regressions, decision tree, random forest, discriminant function analysis, and multilayer perceptron (MLP). To the thyroid diseases classification.

***Keywords*---**thyroid disease, machine learning algorithms, predictive models.

Introduction

Thyroid illness is an endocrinology subgroup that is one among the diseases that are the most misunderstood and undiagnosed [1] [2]. According to the WHO, problem related to Thyroid gland disease is the second most common endocrine ailment in the world, after diabetes. About 2% and 1% of people, respectively, suffer with hyper functioning hyperthyroidism and hypothyroidism. Men's prevalence is around a tenth that of women's. Hyperthyroidism and hypothyroidism can be caused by thyroid gland malfunction, which can be after pituitary gland malfunction or third to hypothalamus dysfunction. In some locations, goitre or functional thyroid nodules can form as a result of a dietary iodine deficiency, with an occurrence of about 15 percent. Endogenous antibodies (autoantibody) can cause havoc in the thyroid gland, and it can also be the site of a number of malignancies [3].

Experts claim that initially illness identification, diagnosis, & treatment, are crucial in reducing the progression of the disease and mortality. For a number of disorders, early diagnosis and possible treatments improves successful therapy. Despite several clinical, trials detection is generally seen as a challenging job [4]. The butterfly-shaped thyroid gland found towards the bottom of the throat. It is made up of levothyroxine (T4), two functional thyroid hormones, and triiodothyroxine (T3), which regulate blood pressure, heart rate and body temperature. Thyroid disease, moreover, is among the most popular disease in the country, and it's usually triggered by iodine shortage, however this can even be triggered by some other factors. Whereas thyroid gland seems to be a hormone-producing endocrine organ that transports hormones all around the body. It is positioned in the front centre of the human body. Hormones of thyroid gland are mainly responsible for metabolism, and also keeping the moisture of body balanced. Thyroid problems are of two different types: hyperthyroidism & hypothyroidism. Mining of data [5] seems to be a technique of searching that is semi-automated relationships in large databases. ML techniques are one among the most powerful approaches to a wide range of complicated problems [6]. Classification seems to be a method of data gathering (machine learning) that is used for detecting & recognising various diseases, which include thyroid cancer, something that we explored and classified here simply as ml algorithms play a very critical part in characterising thyroid disease as they are effective and

successful and aid in classifying [7]. Despite the fact that artificial intelligence &machine learning have been used in medical since its inception [8], recently, there has been a push to address the need for machines to comprehend solution of healthcare. Due to this, experts anticipate that ML would turn into standard healthcare system in the coming years [9]. Thyroid hyperthyroidism occurs when the thyroid gland generates too many hormones related to thyroid. A rise in thyroid hormone levels causes hyperthyroidism [10]. Increased temperature sensitivity, dry skin, weight loss, thinning of hair, elevated heart rate, excessive sweating, high blood pressure (BP), neck expansion, nervousness, shorter menstrual cycles, hands shaking, and abnormal stomach motions are only a few of the symptoms. Hypothyroidisms are disordered in that thyroid gland does not function properly. A decrease in thyroid hormones generation causes hypothyroidism. Whereas Hypo means "insufficient" or "less" not medical jargon. Thyroid gland problem and inflammation are twice major sources of hypothyroidism. Low heart rate, obesity, neck swelling, enlarged temperature sensitivity, hand numbness, hair problems, dry skin ,heavy menstrual cycles, and digestive nuisance are all just a couple of small symptoms. If these symptoms aren't handled, they can get worse with time [12].

Background and related work

Several researches in the literature focus with the detection of thyroid illnesses using hormonal measures and private data from the patient, such as age and sex. Some research, for example, employ machine learning classifications and predictive model, whilst others use deeper CNN architectures. In this study, Thyroid disorder is diagnosed with several classified models based on factors including TSH, T4U, as well as goitre. It argument is supported by a number of grouping algorithms, including K-nearest neighbour [13]. The algorithm Naive Bayes and support vector machine is used. These studies were conducted with the rapid miner tool, and this results show to K-nearest neighbour is extra successful than Naive Bayes in diagnosing thyroid problems. The researchers employed data mining classifiers to diagnose thyroid disease. When diagnosing a condition, thyroid issue is an important element to consider. In this work, Naive Bayes and KNN classifiers were utilised. These two classifiers are comparing with the Rapid miner programme. The K-nearest neighbour classifier with 93.44% accuracy was found to be the most dependable against Naive Bayes classifiers with 22.58% accuracy. The suggested KNN approach results in better outcomes by enhancing accuracy of classification. Due to this, Naive Bayes gives linear, elliptical, or parabolic decision boundary; hence it is significantly advantageous to have decision boundary consistency offered by KNN. Because the components are interrelated, KNN beats most techniques. Thyroid diseases are most common affecting humans. The hypothyroid data utilised in this study has been taken from Irvine's data repository (UCI) of University of California. The decision stump tree approach was proven to be less successful than the J48 technique. Diagnosing disease in the field of medicine is a daunting task. A variety of data mining approaches are employed in the decision-making procedure. We reduced dimensionality to decide selection of characteristics from the earlier findings, and we defined hypothyroidism using J48 and different classification approaches for decision stump data mining. The uncertainty matrix is used to evaluate the precision and error rate of a classifier's output. The J48 Algorithm offers higher

accuracy percentage (99.58%) and lower rate of error, than that of the decision stump tree [14]. Classification is the important data mining approaches under supervised learning and is used to describe preset data sets. The categorization is extensively used in the healthcare industry to help take decision in medical diagnosis as well as supervision. The data for this research came from a renowned laboratory in Kashmir. The ANACONDA3-5.2.0 platform will be used for the whole research study. Different methods of Classification viz. Decision Tree, K closest neighbours, Naive Bayes etc. may be employed in an experimental investigation. With a 98.89 percent accuracy rate, the Judgment Tree is the most accurate of the other classes [15]. Thyroid disease is a long-term condition that affects the population globally. In healthcare, data mining proves to yield outstanding outcomes in the prediction (prognosis) of multiple diseases. Data mining approaches for prediction have a high level of accuracy and a cheap cost of prediction. Another key advantage is that prediction takes only a short amount of time. In this study, I analysed thyroid data using classification algorithms and came up with a conclusion. The effectiveness of a model is essentially governed by two variables. Forecast precision is the first, while prediction time is the second. According to the data, Nave Bayes forecasting took only 0.04 sec. It is less precise than Random Forest & J48. The Random Forest model has a prediction accuracy of 99.3 percent when we looked at it. On the other hand, the model's creation time is higher than further two iterations. As a result, for hypothyroid prediction J48 is the best yielding model having 99% accuracy and its execution time (0.2 sec) is quite less than Random Forest [16]. The goal of this job is to offer a plan based on data mining for improving precision of hypothyroidism diagnostic by combining queries of patients and test findings throughout the diagnostic procedure. Another objective is lowering the hazards associated with interventional studies of dialysis. The conclusive outcome whether new samples are hypothyroid, was inferred using statistics from machine learning repository of UCI, providing 3165 sample data (155 of hyperthyroid and the rest hypothyroid). To eliminate the imbalanced distribution, several sampling approaches were applied in the data gathering, and hypothyroidism models were constructed using K Nearest Neighbour, Logistic Regression & SVM classifiers. In this context, the hypothesis established the influence of sample methodologies on analysis of hypothyroidism. The Logistic Regression classifiers delivered the best outcomes amongst the entire model constructed. For this study, performance of model trained with dataset uses over-sampling approaches is found as: Accuracy - 97.8%, F-Score - 82.26%, RoC - 93.2%, and Matthews Correlation Coefficient - 81.8% [17]. The goal of this study is to develop a machine learning system using Random Forest method to detect diabetes inside patient near the beginning and correctly. Random Forest methods are used for classification phase and regression phase of application. The performance ratio is greater when compared to other methods. The proposed model produces the best prediction results for diabetes diagnosis, and the findings show that the system is accurate, effective and, quick in diagnosing the diabetes. [18]. Breast cancers are second more frequent cancer in females after all other malignancies. The purpose of the research article is to provide diagnosis of breast using avant-garde methods. Incorporating recent scientific developments into breast cancer survival prediction models. To build prediction models, we employed a large dataset and three standard data mining methods. We utilised 10-folded cross-validation procedures to test the impartial estimate of all the predictive models in subject for accuracy

evaluation. According to the findings, the Bay is safe area to visit. The RBF Networks are the next-best predictors by 93.44% accuracies on the propose sample, followed by the Nave Bayes with 97.36% on the holdout sample. We used two criteria in this study to assess varying breast cancer survivorship prediction model, whether benign or malignant [19]. The more recent study focus on the categorization of thyroid-illness in two most common thyroid dysfunctions in the common people. Radial Basis Function Networks, Naive Bayes, Multilayer Perceptrons, Decision Trees, were four classification models, researchers looked at and compared. The results show that all of the above-mentioned classification models are very accurate, with the Decision Tree model yielding uppermost categorization of score. Data from UCI machine learning repository and Romanian data website were used to build and evaluate the classifier [20]. KNIME Analytics Platform along with Weka is two datasets. The categorization models were developed and tested using data mining techniques as the base. With literature survey, a range of learning in the subject of categorization of thyroid employs different data mining approaches to build strong classifiers. The authors examined how four classification models may be used on thyroid data to assist categorise thyroid dysfunctions including hyperthyroidism along with hypothyroidism. The decision tree model was proper categorization model in all scenarios that were examined.

Proposed approach

Data Collection

Since they currently occupy a major relevance in the healthcare area as well as assist us in identification and classification of diseases, methods are utilised in the rapid and effective cure of thyroid problems as well as many other diseases. As a consequence, we became able to gather a large amount of information on thyroid illnesses, which we used in our disease identification studies. The data was taken from Iraqi citizens, and Thyroid illness was the focus of the material collected. Data was taken on around 1250 different people, both females and males, age group ranging from one year to one year. Patients having hypothyroidism and hyperthyroidism, and also healthy adults without thyroid disease, were included in these samples. The information was gathered over duration from one - four months with the goal of using machine learning techniques to classify thyroid diseases. Age, Gender, T3 (triiodothyronine), TSH (thyroid stimulating hormone), T4 (thyroid hormone), and some other characteristics are one among the data gathered. Because the data collected included 17 variables or traits, all of which have been included in our research comprising (id, gender, age, ill, query hypothyroid, thyroid surgery, pregnant, query hyperthyroid, etc).

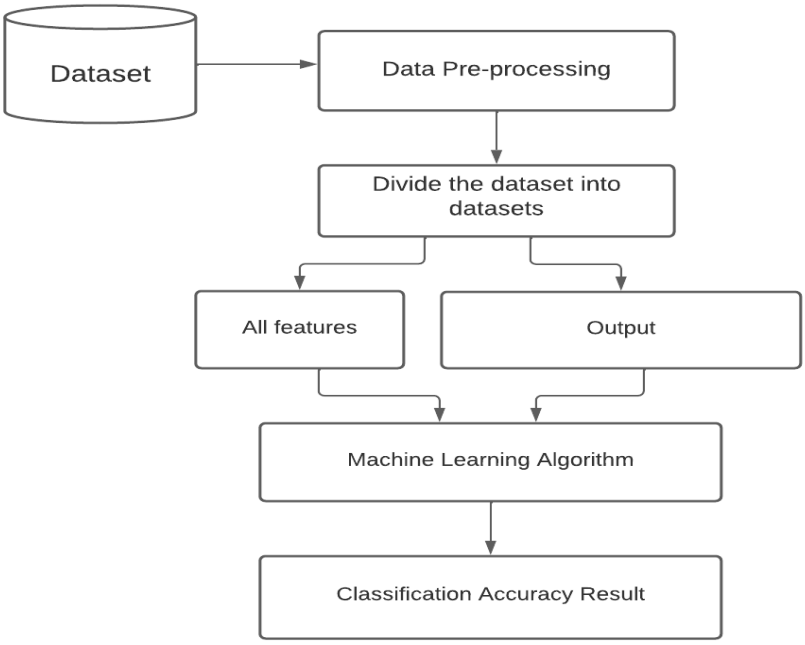

Fig. 1 Representing Architecture of data processing

Data Pre-processing

The pre-processing of statistics is highly significant and crucial stage in taking out of data, as it has a positive impact on the data, as it is uses to disclose the information via analysing present data and uncovering the missing data, since it carefully studies the data. Cleaning data, prepping the data, and so on are all part of the pre-processing process. Inside this stage or step, we cleaned and organised the records that if we have capable to obtain, within which we recognised a set of incomplete information in this statistics where the lacking characteristics were recognised, and between these missing characteristics be T4 by quantity 155 and T3 by amount 115, in which we have capable to Interpret this deleted information by replace this by the valuation of intermediary, as well as after worked in this manner, we have capable to acquire the information in a higher as well as positive way and liberated of missing data. With the MLP method, we also applied normalising techniques.

Data Machine Learning Techniques

Machine learning algorithms are being used to distinguish between three types of thyroid illness. Hyperthyroidism is the very first, hypothyroidism is the second, and stable people with no thyroid disorders are the third.

Support Vector Machines

SVM algorithm is machine learning based algorithm which is associates with supervised learning and data mining approach that may be used to find the best predictors of energy usage. To address our query, the researchers employed prominent categorization approaches such as best boosting trees, subset selection, and generalised additive model. To choose a subset of predictor which more significantly predicted consuming with a linear association, we used backward, forward and optimal subset selection. The SVM proposed a method that used iterative binary splitting to categorise the predictor spaces into sample areas using a tree-based algorithm. The boosting tree approach was chosen since it is regarded to be once among the mostly used tree-based algorithm solutions. SVM is too effective at dealing with data with a lot of dimensions.

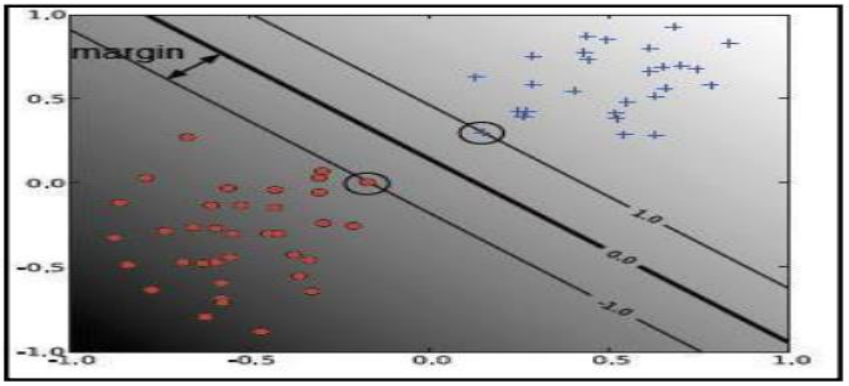

Fig. 2 Support Vector Machine Classifiers

Random forest

The random forest calculates the average response of each energy consumption predictor. The relative distance between every response was in the mean of every predictor is then added by a random forest for merely a total amount of distance between each response would be from the database's averages from each sample. Those who were regularly distant from the response variable in each sample will have a high maximum distance. A programme that determined the mode of each answer was used to detect rates that repeatedly categorise the data. The research labelled an answer as possibly high in power consumption if the style of response accounted for more than 90% of the overall number of questions. There are a lot of replies that have been noted. A visual analysis of these replies revealed that the subjects had all selected same response.

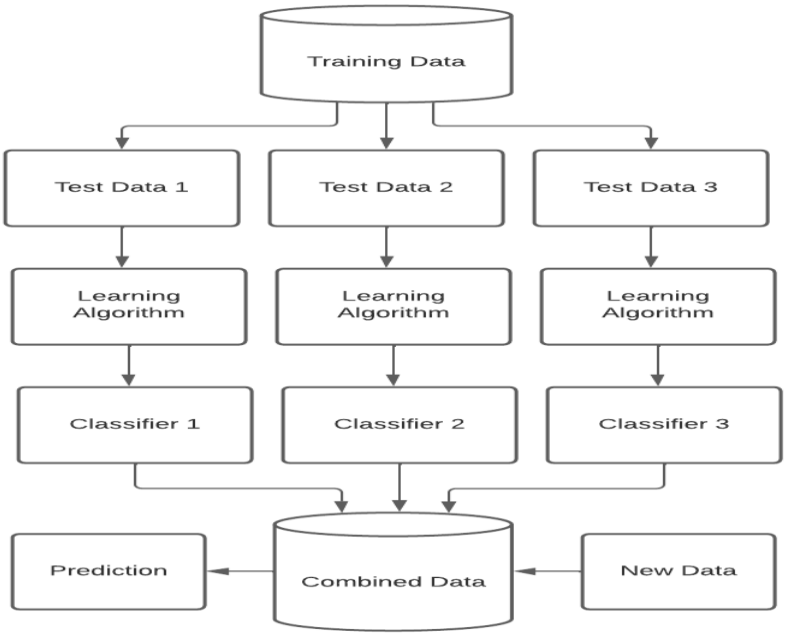

Fig. 3 Flowchart of RF classification

Decision Tree

Decision tree based algorithm is relies on the decision-boosting system that evaluated for forecasting energy usage component, with the purpose of finding the most significant predictors of consumption using a tree-based method. I utilised the decision tree technique to accomplish this. Hundreds of trees should be fitted in order for the product to develop over time, which are grown using knowledge from the past tree. Tuning parameters include shrinkage parameter, the quantity of trees, and quantity of breaks in every tree.

Naïve Bayes

Using Nave Bayesian, several generalized additive models incorporating subset selection independent variables, variables having the biggest relative impact on classification, and a mixture of different factors from both can be compared. It directly compared each one of the best models' prediction accuracies. By fit naive Bayes through varying combination of splines, two degree polynomials, and sequential predictor's variable, it limited correlations among individual predictor's and responses. To construct predictors with nonlinear links to our response variable, we use much more polynomials and splines.

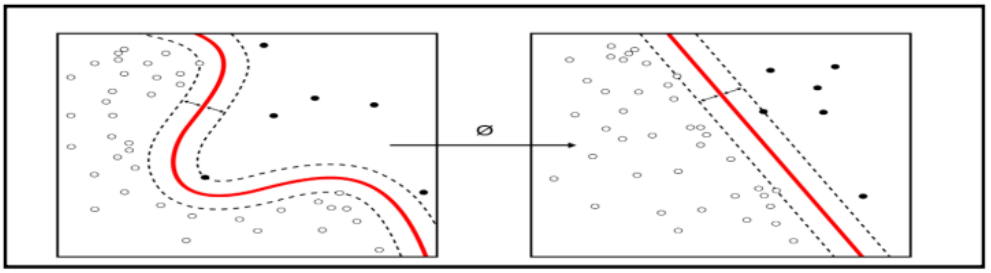

Fig. 4 Naïve Bayes representation

Logistic Regression

Logistic regression, which is included in the Supervised Learning methods, is widely uses in Machine Learning (ML) algorithms. These methods are used to estimate a dependent categorical variable with a group of independents factors. LR is used to predict overall contribution of a dependent categorical variable. As a consequence, the value obtained must be categorical or solitary. It may be false or true, 0 or 1, No or Yes, and so on, rather than precise figures like 0 and 1, it returns probabilistic values in the middle. In terms of application base, Logistic Regression is more similar to Linear Regression. And in this direction to solve regression problems, linear regression is employed, whereas logistic regression is being utilize to solve classification problems [25].

KNN

The k-nearest neighbour techniques, unlike earlier approaches, use the information directly for categorization instead of first creating the models [26]. Such type output, no extra model building is essential, and model's single component is k, the quantity of nearest neighbours that can be used in calculating class label: the result of p(y/x) is basically ratio of users of elegant between some of the k^{th} nearest neighbours of x. This model's stability is affected by varying values of k. One of the key advantages of KNN over all various algorithms is no difficulty of use. Neighbours may be able to explain the classification outcome. This case-based reasoning might be useful in circumstances when black-box models are insufficient. The determination of the case neighbourhood, which necessitates establishing a metrics which estimates the total distance linking different data items, is the main drawback of k-nearest neighbours [27].

Multi-Layer Perceptron's

A feed forward ANN that creates sequence of outcomes from the collection of inputs is called multi-layer perceptron's. Between the input and output layers of an MLP, numerous layer of inputs node create a directed graphs. Back propagation is used by MLP to trains the networks. MLP are a method of DL. The multilayer perceptron's are neural network which links many layers in the directed graphs, with the signal travelling exclusively in single direction among nodes. Except for the inputs node, each node has non-linear activation functions. Back propagation is a supervised learning mechanism uses by MLPs. MLP is DL

based approach that makes use of many layers of neurons. There are few issues in the supervised learning, which are parallel distributed processing analysis another computational biology, MLP is a widely utilised approach. Speech recognition, picture recognition, and automated translation are some of the applications [28].

Linear Analysis

It's one among the mostly popular methods for reducing dimensionality. Among other things, it's uses in the pattern classification and machine learning techniques. LDA is a method of projecting things from 3D space to 2D space. It's done to avoid frequent dimensional problems while also minimising spatial and operating expenditures. Discriminant function analysis, a classifier approach, is uses to create machine learning method. Framework for reducing dimensionality are utilised in the variety of applications, including predictions and picture identification [29].

Result analysis

The proposed research used a large variety of machine learning (ML) algorithms to analyse our data (Random Forest, Decision Tree, Logistic Regression, SVM, Naive Baye, Linear Analysis, KNN, and Multi-Layer Perceptron). Because this is first time we've used this data, we firstly divided dataset into two datasets: out of which 70 percent for training and rest 30 percent for testing. In the initial stage, we applied the attributes in our data to the collection of mechanics listed in the below tables, and outcomes displayed following the application process. This practical section was developed in Python and is regarded a comprehensive and various integrated systems. All sixteen entries are considered as inputs and one output have been get hold into account.

Table 1. Representing the all other algorithms on those classification models

S. No	Apply Algorithms	Accuracy
1	MLP	97.7
2	DT	98.54
3	RF	98.98
4	SVM	92.32
5	KNN	91.25
6	LR	91.48
7	NB	82.12

After deleting the three features, the Naive Bayes method appears to have a higher accuracy of about 90.67, whereas the logistic regression technique, SVM algorithm, and K-Neighbours Classifier algorithm have risen greatly in accuracy whereas the accuracy value of the remaining algorithms has declined. As experience has demonstrated, different accuracy value of the used algorithms employed on above dataset changes as the qualities used throughout the data

vary. When three of the features were removed from the algorithms, the accuracy of certain algorithms declined while the accuracy value of others augmented.

Conclusion

This disease is most common disorders afflicting the entire nation peoples, and the amount of the cases is rising. Our research focuses on the categorization of the thyroid disease into hypothyroidism and hyperthyroidism in light of medical details that demonstrate major abnormalities in thyroid disorders. Algorithms were used to classify this illness. Machine learning produced good results by combining various techniques and constructing two models. The accuracy value of the random-forest method was 98.97 percent that was the greatest accuracy value between all different algorithms tested, in the first model, which had 16 inputs and one output. Based on a prior research, the following qualities were eliminated in the second incarnation. 1- query thyroxine was one of the properties that was eliminated. 2-query hyperthyroid 3-query hypothyroid. We've contained both the enhanced accuracy value of certain algorithms and the control of the accuracy value of others in this section. The accuracies of the Naive Bayes machine learning algorithms were found to rise by 90.67 percent. The MLP algorithm has the maximum precision of 96.4.

References

1. AKGÜL, Göksu, et al. "Hipotiroidi Hastalığı Teşhisinde Sınıflandırma Algoritmalarının Kullanımı." Bilişim Teknolojileri Dergisi 13.3 (2020): 255-268.
2. Aswad, Salma Abdullah, and Emrullah Sonuç. "Classification of VPN Network Traffic Flow UsingTime Related Features on Apache Spark." 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE, 2020.
3. Aytürk Keleş and Keleş, Ali. "ESTDD: Expert system for thyroid diseases diagnosis." International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 11 | Nov -2017 34.1 (2017):242- 246
4. Azar, a.T, Hassanien, A.E. and Kim, T. Expert system based on neural fuzzy rules for thyroid disease diagnosis, Computer Science, Artificial Intelligence, arXiv:1403.0522, Pp. 1-12,2012.
5. Banu, G. Rasitha. "A Role of decision Tree classification data Mining Technique in Diagnosing Thyroiddisease." International Journal of Computer Sciences and Engineering 4.11 (2016): 64-70.
6. Banu, G. Rasitha. "A Role of decision Tree classification data Mining Technique in Diagnosing Thyroid disease." International Journal of Computer Sciences and Engineering 4.11 (2016): 64-70.
7. Begum, Amina, and A. Parkavi. "Prediction of thyroid disease using data mining techniques." 20195th International Conference on Advanced Computing & Communication Systems (ICACCS). IEEE, 2019.
8. Bolotnikova, Anastasia, Hasan Demirel, and Gholamreza Anbarjafari. "Real-time ensemble based face recognition system for NAO humanoids using local binary pattern." *Analog Integrated Circuits and Signal Processing* 92, no. 3 (2017): 467-475.

9. C. Fan, F. Xiao, Z. Li, J. Wang. Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review. Energy Build. 2018, 159, 296–308.
10. Chandel, Khushboo, et al. "A comparative study on thyroid disease detection using K-nearest neighbour and Naive Bayes classification techniques." CSI transactions on ICT 4.2-4 (2016): 313-319.
11. Chandio, Jamil Ahmed, et al. "TDV: Intelligent system for thyroid disease visualization." 2016International Conference on Computing, Electronic and Electrical Engineering (ICE Cube). IEEE, 2016.
12. Chaurasia, Vikas, Saurabh Pal, and B. B. Tiwari. "Prediction of benign and malignant breast cancerusing data mining techniques." Journal of Algorithms & Computational Technology 12.2 (2018): 119-126.
13. D. Mora, G. Fajilla, M. Austin, D. Simone. Occupancy patterns obtained by heuristic approaches: Cluster analysis and logical flowcharts. A case study in a university office. Energy Build. 2019, 186, 147–168
14. Dasarathy B. Nearest neighbor pattern classification techniques. Silver Spring, MD: IEEE Computer Society Press; 1991.
15. Dr. Srinivasan B, Pavya K "Diagnosis of Thyroid Disease: A Study" International Research Journal of Engineering and Technology Volume: 03 Issue: 11 | Nov – 2016
16. Dreiseitl, Stephan, and Lucila Ohno-Machado. "Logistic regression and artificial neural network classification models: a methodology review." Journal of biomedical informatics 35.5-6 (2002): 352-359.
17. Gupta, Puneet, Vijay Kumar Sharma, Naman Mittal, Raghav Bansal, and Himanshu Gupta. "Ai Enabled Virtual Environment Simulator." (2020).
18. Heuck, "World Health Organization," 2000. [Online]. Available: https://www.who.int/.
19. Iqbal, Md, Vimal Kumar, and Vijay Kumar Sharma. "Krishi Portal: Web Based Farmer Help Assistance." International Journal of Advanced Science and Technology 29 (2007): 4783-4786.
20. Keles, A. ESTDD: Expert system for thyroid diseases diagnosis, Expert Syst Appl., Vol. 34, No.1,Pp.242–246, 2008.
21. Khushboo Taneja, Parveen Sehgal, Prerana "Predictive Data Mining for Diagnosis of Thyroid Disease using Neural Network" International Journal of Research in Management, Science & Technology (E-ISSN:2321- 3264) Vol. 3, No. 2, April 2016
22. Kouroua, K., Exarchosa, T.P. Exarchosa, K.P., Karamouzisc, M.V. andFotiadisa, D.I. (2015) Machinelearning applications in cancer prognosis and prediction, Computational and Structural BiotechnologyJournal, Vol. 13, Pp.8–17.
23. Ripley B. Pattern recognition and neural networks. Cambridge: Cambridge University Press; 1996.
24. Sharma, Vijay Kumar, Vimal Kumar, Md Iqbal, Sachin Tawara, and Vishal Jayaswal. "Virtual Mouse Control Using Hand Class Gesture."
25. Shukla, A. & Kaur, P. (2009). Diagnosis of thyroid disorders using artificial neural networks, IEEE International Advance computing Conference (IACC 2009)– Patiala, India, pp 1016-1020.
26. Sindhya, Mrs K. "EFFECTIVE PREDICTION OF HYPOTHYROID USING VARIOUS DATAMINING TECHNIQUES."

27. Tiwari, Durgesh, and Vijay Kumar Sharma. "A Review on Conventional and Lightweight Security Techniques in Mobile and IoT Devices."
28. Travis B Murdoch and Allan S Detsky. The inevitable application of big data to health care. Jama,309(13):1351–1352, 2013.
29. Tyagi, Vidhi, Shivam Arora, Sattyam Gupta, Vijay Kr Sharma, and Vimal Kumar. "Architecture of an IoT-based Women Safety System." *Architecture* 29, no. 5 (2020): 3670-3676.
30. Umar Sidiq, Dr, Syed Mutahar Aaqib, and Rafi Ahmad Khan. "Diagnosis of various thyroid ailments using data mining classification techniques." Int J Sci Res Coput Sci Inf Technol 5 (2019): 131-6.
31. V. Cerqueira, L. Torgo, M. Mozetic. Evaluating time series forecasting models: An empirical study on performance estimation methods. Mach. Learn. 2020, 109, 1997–2028.
32. VijiyaKumar, K., et al. "Random Forest Algorithm for the Prediction of Diabetes." 2019 IEEEInternational Conference on System, Computation, Automation and Networking (ICSCAN). IEEE, 2019.
33. W. Kleiminger, C. Beckel, T. Staake, S. Santini. Occupancy Detection from Electricity Consumption Data. In Proceedings of the 5th ACM Workshop on Embedded Systems for Energy-Efficient Buildings, Rome, Italy, 14–15 November 2013; pp. 1–8.