International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

DETECTION OF THYROID DISEASE USING MACHINE LEARNING ALGORITHM

Ms.M.ANITHA[1], Mr. EJJIVARAPU NAGARAJU [2], Ms. GUNAGANTI.GOWTHAMI [3]

#1 Assistant professor in the Masters of Computer Applications in the SRK Institute of Technology, Enikepadu, Vijayawada, NTR District

#2 Assistant professor in the Masters of Computer Applications in the SRK Institute of Technology, Enikepadu, Vijayawada, NTR District

#3 MCA student in the Masters of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District

ABSTRACT: Thyroid gland diseases are a critical aspect in medical diagnosis and prognosis, which is a tough concept to grasp in the medical profession. The thyroid gland is one of the most important elements of the human body. Thyroid hormones play an important role in metabolic regulation. Thyroid hormones play a role in the body's ability to regulate metabolism, and both excess and deficiency can be harmful. Machine learning is used to predict illness and to investigate thyroid disease categorization models using hospital datasets. For dealing with dynamic learning tasks like medical diagnosis and prediction, a robust knowledge base constructed and deployed as a hybrid model is needed. It may be able to detect and decrease thyroid activity using basic machine learning algorithms. A support vector machine (SVM) model is commonly used to forecast the likelihood of a thyroid patient. When a patient is at risk for thyroid disease, our system must provide recommendations such as home remedies, warnings, precautions, medicines, and so on.

1. INTRODUCTION

Some of the most cutting- edge operations of machine biology are in medical care. Gathering information for medical complaint prophecy was essential. various intelligent prophecy algorithms are used to find problems with a product or service beforehand on. Although the Medical Information System excels at handling data sets, there are presently no intelligent technologies available for furnishing a prompt prognostic of cases' affections. In the end, machine knowledge algorithms play a significant part in diving delicate andnon-linear challenges throughout the development of the vaccination model. Any complaint prophecy model must bear naming rates from various data sets that may be employed as description in a healthy situation as precisely as doable. A misclassification may lead to a happy case being placed in a bad care setting. It's also of the utmost cardinal significance that any possibility of vaccinating against thyroid complaint be considered. The thyroid gland is a gastrointestinal endocrine gland. It's erected in the mortal neck's lower region, under the Adam's apple, and helps the body store thyroid hormones, which in turn impacts the body's rudimentary

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

metabolic rate and its capability to produce proteins. These hormones calculate on the rate at which the heart beats and the calories are burned to regulate the body's metabolism. Thyroid hormones play a part in regulating the body's metabolism through their chemical makeup. These organs produce the hormones thyroxine(docked to T4) and triiodothyronine, which are responsible for regulating the body's metabolism(docked T3). Thyroid hormones have a pivotal part in artificial processes, as well as in the structure and conservation of structures and the regulation of body temperature. Two actuated thyroid hormones, T4 and T3, make up the thyroid glands. These hormones have a pivotal part in regulating proteins, maintaining internal body temperature, transporting energy throughout the body, and promoting cell proliferation. Iodine, along with T3 and T4 harmones, is a abecedarian structure block of the thyroid glands and is only inactive in a sprinkle of extremely rare yet critical situations. Both hypothyroidism and hyperthyroidism can affect from inadequate situations of these hormones in the body. There's a wide range of possible causes for both hyperactive and underactive thyroid. Several pharmaceutical options live. Iodination insufficiency, thyroid atrophy, and a lack of thyroid hormone- producing enzyme are all side goods of thyroid surgery.

2. LITERATURE REVIEW

Machine learning and deep learning techniques have been used in several research studies for thyroid disease prediction due to recent technological advancements in data processing and computation. Early detection and classification of this disease as cancer, hypothyroidism, or hyperthyroidism is beneficial for timely treatment and recovery. The survey of the literature is carried out using peer-reviewed article databases such as Google Scholar and Scopus. To find the most current papers in our analysis, we conducted searches over the last five years. The relevant articles were chosen using the keyword combinations "Thyroid sickness," "Thyroid cancer," "machine learning," and "deep learning." We refined the search terms and utilised a tight keyword search because the number of retrieved results is considerably higher for locating relevant articles. During our initial screening, we discovered over 100 relevant articles. We further evaluated those publications and identified 25 that are highly related to our work. Machine learning and deep learning algorithms are utilised to diagnose thyroid illness and thyroid cancer. Because the application of these strategies differs for each task, they are separated.

3. IMPLEMENTATION

Data from the UCI machine learning archives is used in both the research papers and the model classifications used in the prediction of thyroid disease. A good knowledge base that can be centralised and used as a hybrid paradigm must be maintained in order to address complex learning concerns like medical diagnostics and statistical tasks. Additionally, we provided a selection of machine learning and thyroid diagnostic

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

approaches. Thyroid disease risk was estimated using machine learning algorithms like the Vector Support Machine.

In this project, we are using a support vector machine learning algorithm called SVM to determine if a patient's reported data is normal or indicative of a higher risk for thyroid disease; if the latter is the case, the application will present the patient with appropriate dietary and prescription guidelines. In this study, we are developing a prediction model by training a support vector machine using the UCI machine learning THYROID illness dataset. The purpose of applying new patient test data to a trained SVM model is to predict whether a patient is normal or at risk of thyroid disease.

4.DATASET

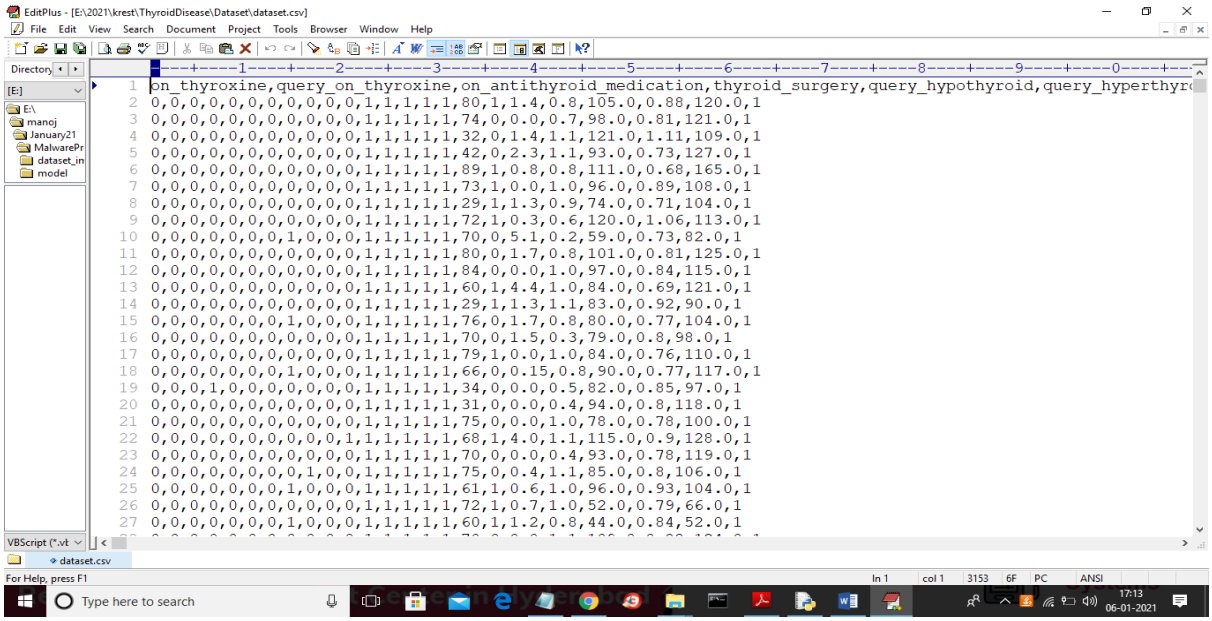
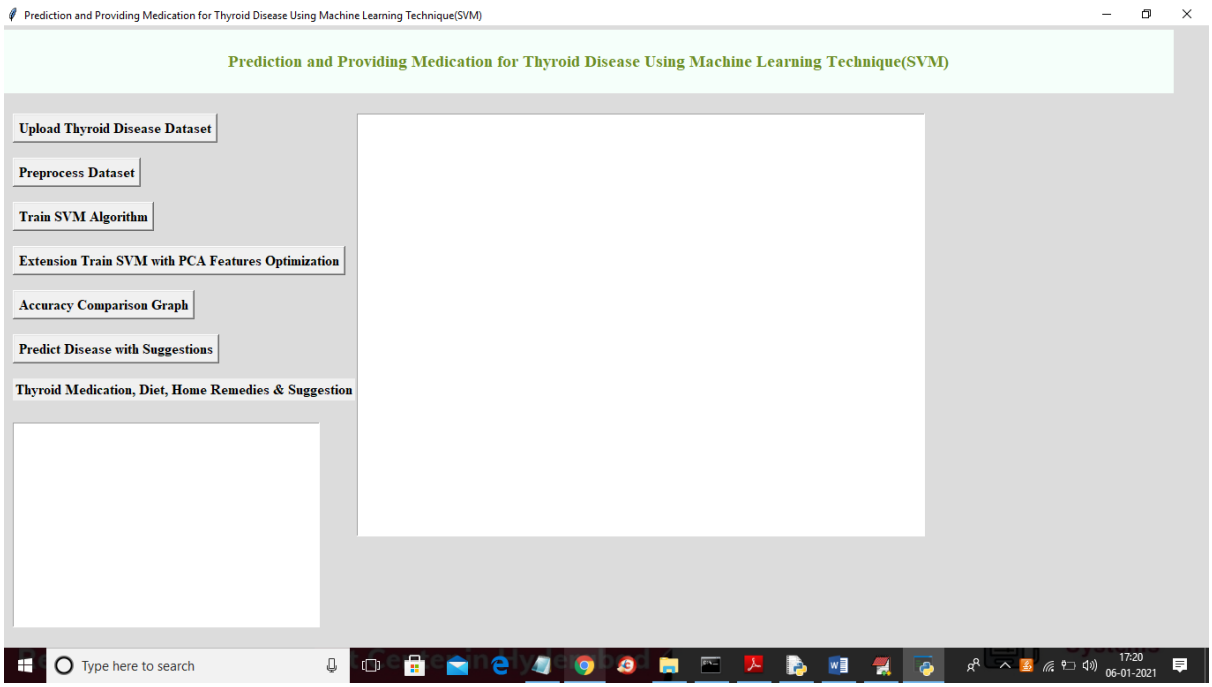


Fig 4:Data Set Values

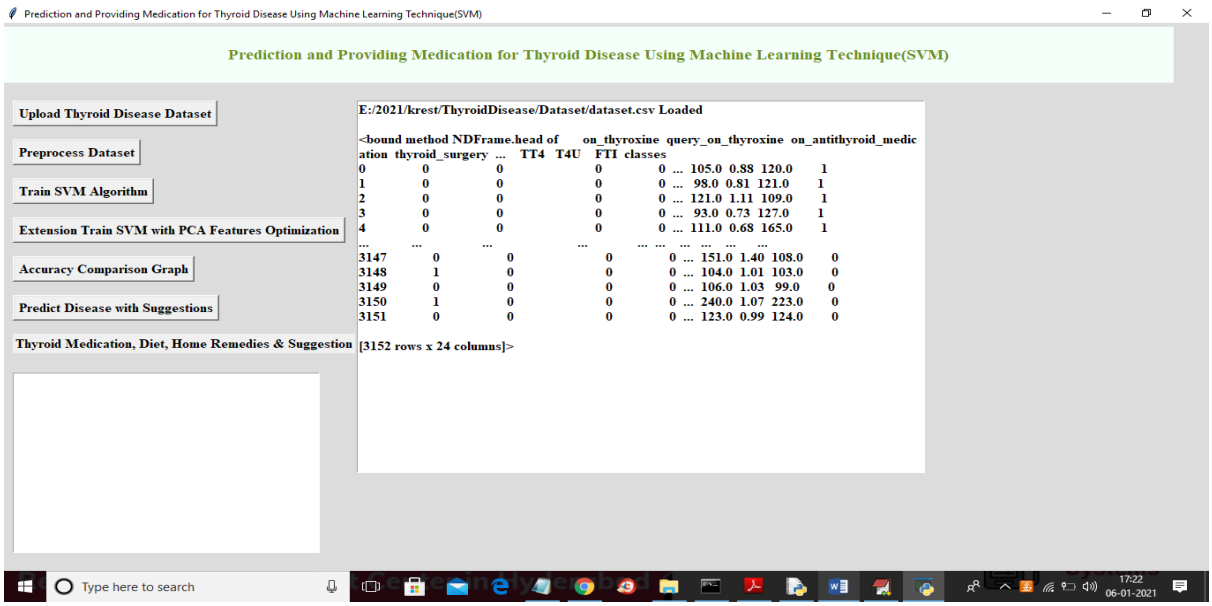
In above dataset first row contains column names and other rows contains values as 0 or 1 and if patient is under thyroid medication or surgery then its column value will be 1 else 0 and in last column contains class label as 0 or 1 where 0 means patient record is normal and 1 means patient record contains thyroid disease. In this dataset more than 3000 rows are there and 24 columns are available. For prediction all 24 columns are not available so we are applying PCA (principal component analysis) feature selection algorithm as extension concept to optimize features or to reduce columns or features which are not important for prediction. PCA will remove unnecessary columns from the dataset and use only important attributes to train SVM algorithm and due to optimize features SVM prediction accuracy can be increase.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

5. EXPERIMENTAL RESULTS

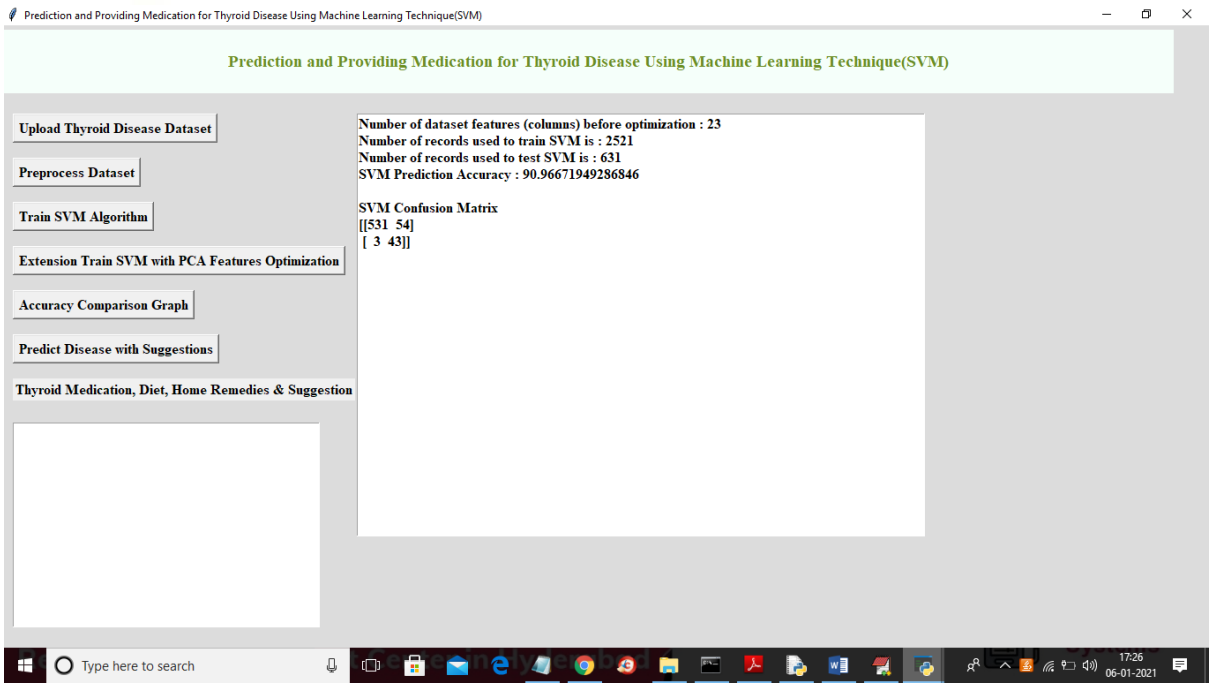


To upload a dataset related to thyroid disease, the aforementioned screen must be navigated to before the screen shown below can be accessed.

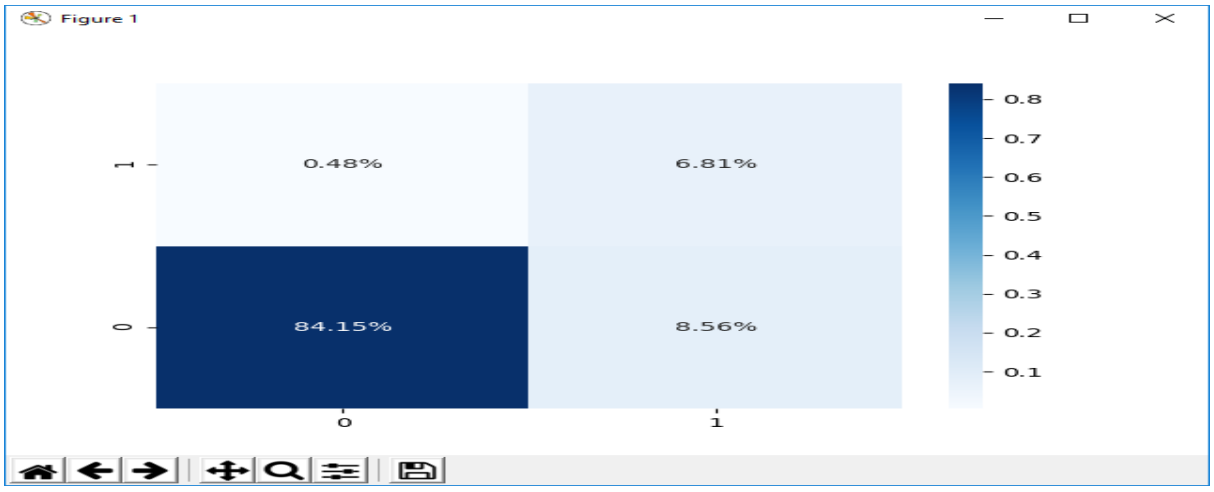


You can remove missing and NAN values from the dataset and split it into X and Y values, where X contains all dataset values and Y contains class label value, by clicking the corresponding button in the screen shot above, after which the dataset will be loaded and a few records from it will be displayed..

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

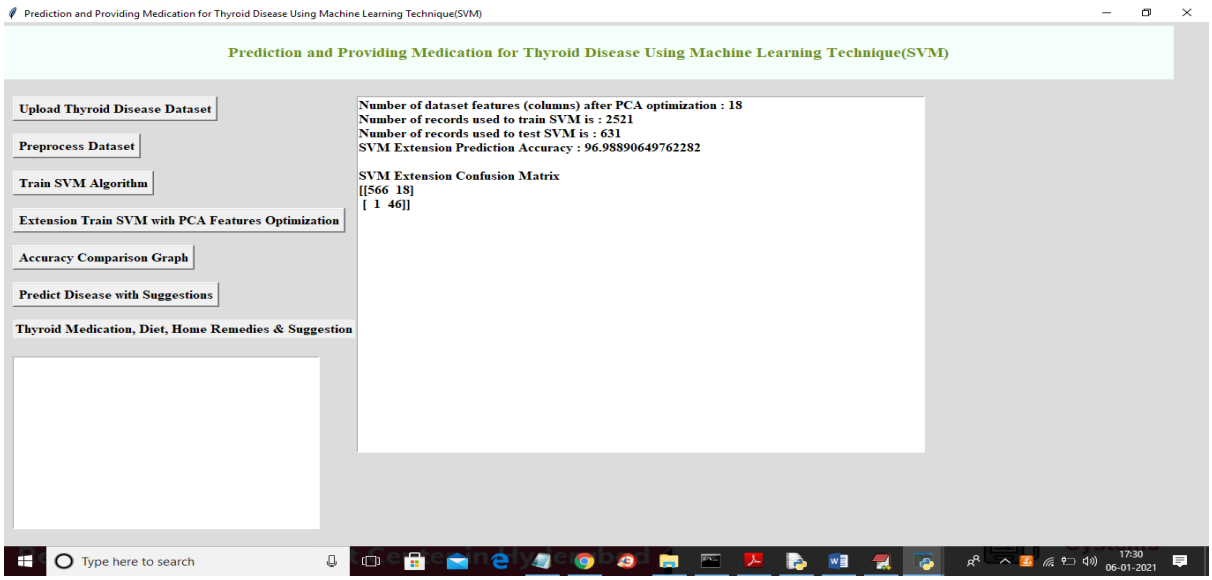


Above, we can see that the dataset has a total of 23 columns, that the SVM algorithm was trained on 2521 records using 631 test records, and that the prediction accuracy was 90.96% using a standard SVM. Furthermore, the application displays a confusion matrix of true and false prediction values, with 531 and 3 representing the correct prediction and 54 and 43 representing the incorrect one, respectively, and a graph representation of the matrix being provided below.

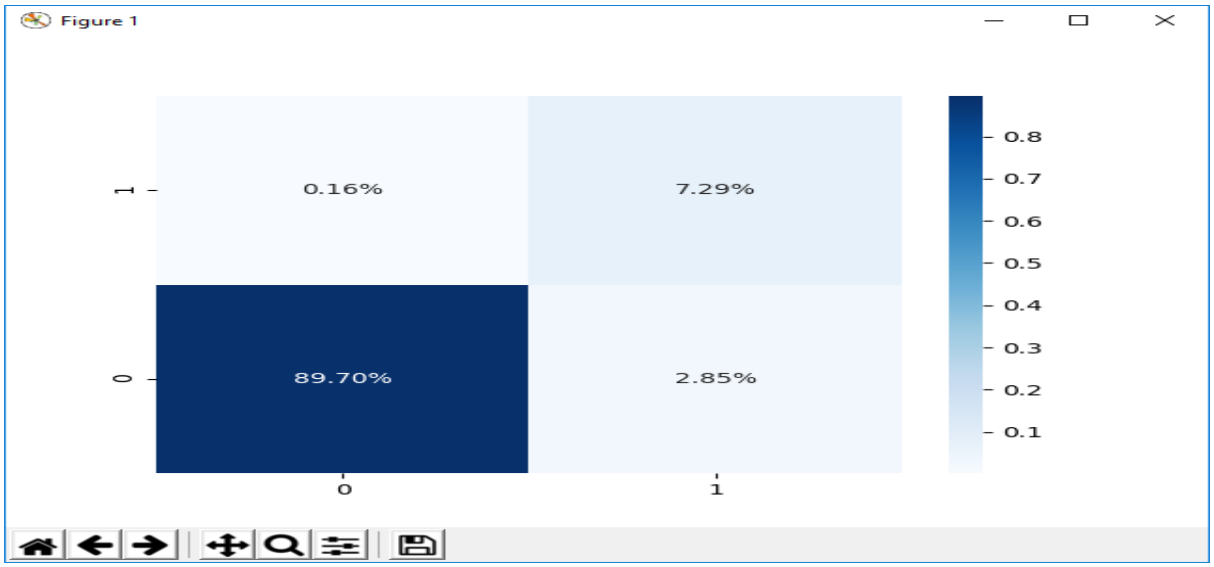


Click the "Extension Train SVM with PCA Features Optimization" button to train the SVM with PCA features optimization and obtain the prediction accuracy shown below the graph (84.15% and 6.81%).

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362



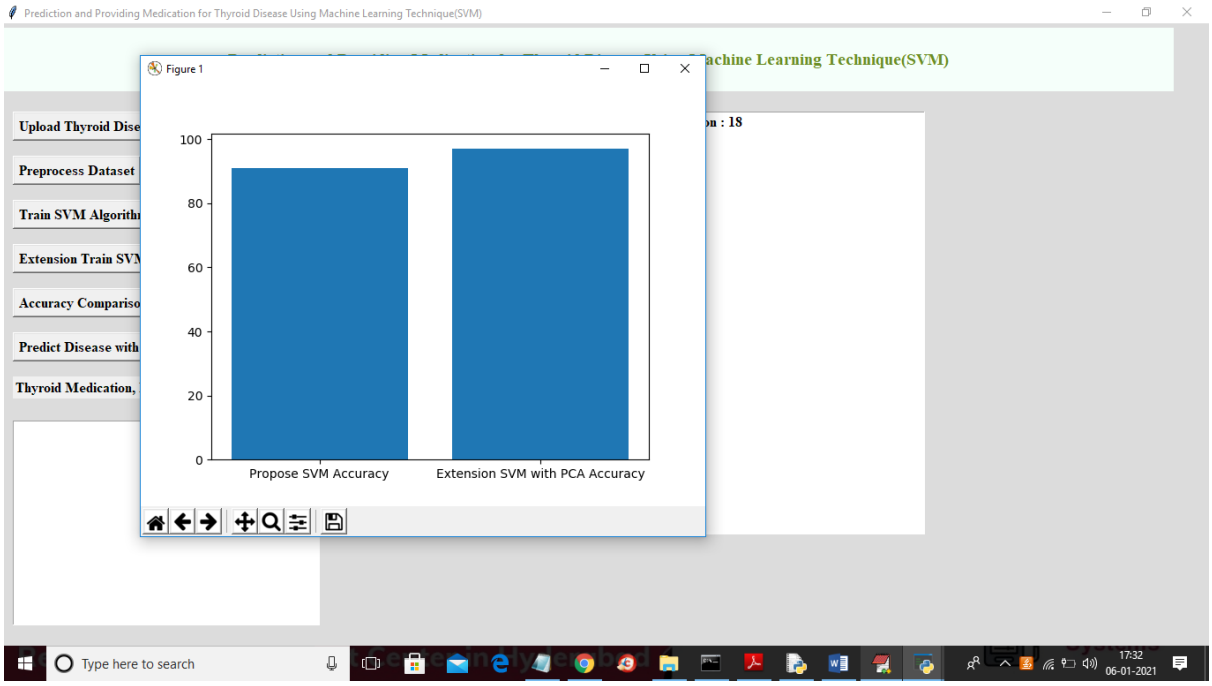
The PCA extension to the SVM shown above achieved a prediction accuracy of 96.08%, and the confusion matrix values it produced were significantly lower than those of standard SVM.



The correct prediction in the above graph is 89.70, and the other values are false predictions. To view the below accuracy comparison graph, click the 'Accuracy

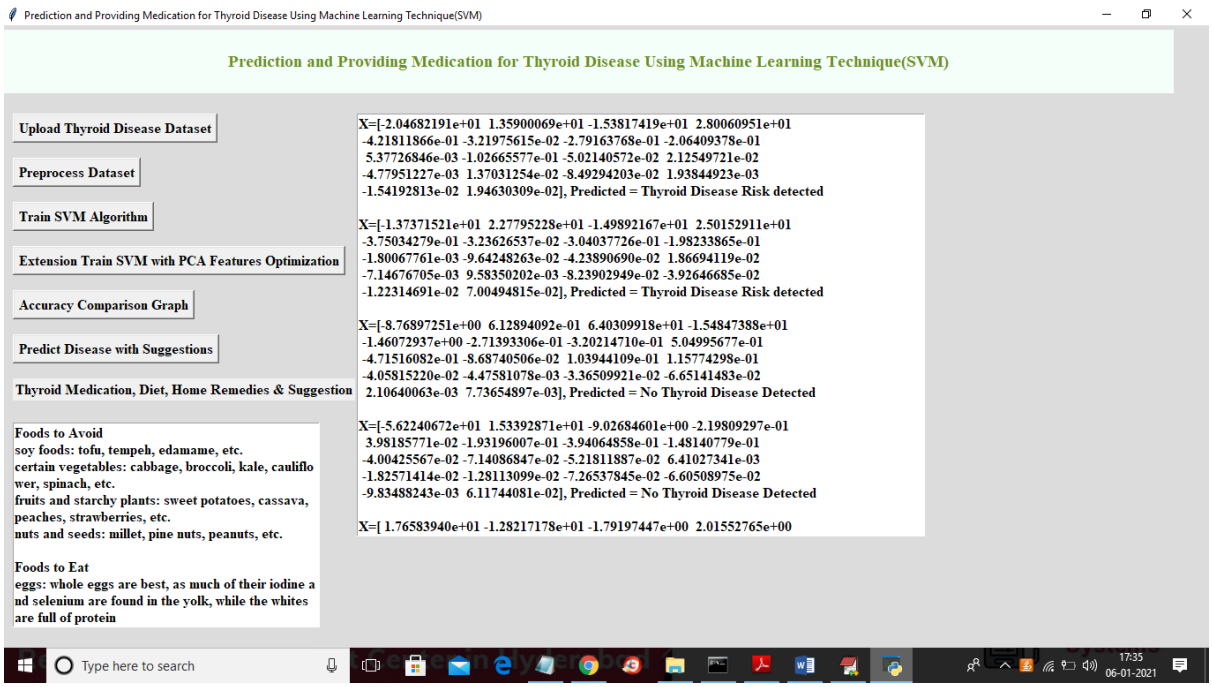
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Comparison Graph' button.



To upload new test data and predict if the data contains thyroid or not, click on the "Predict Disease with Suggestions" button; the x-axis of the above graph represents the name of the algorithm, and the y-axis represents the accuracy of the algorithm. Extension SVM with PCA is superior to normal SVM, as shown in the above graph.

.



Each record's test value is displayed in brackets above; following the brackets is information about whether or not a thyroid risk has been detected, and if one has been, a suggested diet and medication schedule appear in the left box.

6. CONCLUSION

Furthermore, this research studies unique machine learning algorithms for detecting thyroid problems. Many convenient analyses have been developed and used in recent years to correctly and expertly identify thyroid illness. According to the research, the two articles use different technologies with various degrees of accuracy. The bulk of study articles conclude that neural networks outperform other methods. While there is no doubt that medical professionals around the world have made tremendous advances in their ability to identify thyroid issues, it is recommended that people use a smaller subset of the available diagnostic criteria. Having more distinguishing characteristics necessitates more extensive, time-consuming, and costly health assessments.

REFERENCES

[1]Ankita Tyagi and Ritika Mehra. (2018).“Interactive Thyroid Disease Prediction System using Machine Learning Techniques” published on ResearchGate.

[2] YongFeng Wang,(2020). “Comparison Study of Radiomics and Deep-Learning Based Methods for Thyroid Nodules Classification using Ultrasound Images” published on IEEEAccess.

[3] Sunila Godara,(2018). “Prediction of Thyroid Disease Using Machine Learning Techniques” published on IJEE.

[4] Hitesh Garg,(2013). “Segmentation of Thyroid Gland in Ultrasound image using Neural Network” published on IEEE.

[5] L. Ozyılmaz and T. Yıldırım,(2002). “Diagnosis of thyroid disease using artificial neural network methods,” in: Proceedings of ICONIP’02 9th international conference on neural information processing (Singapore: Orchid Country Club, pp. 2033–2036).

[6] K. Polat, S. Sahan and S. Gunes,(2007) “A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted preprocessing for thyroid disease diagnosis,” Expert Systems with Applications,(vol. 32, pp. 1141-1147).

[7] F. Saiti, A. A. Naini, M. A. Shoorehdeli, and M. Teshnehlab,(2009) “Thyroid Disease Diagnosis Based on Genetic Algorithms Using PNN and SVM,” in 3rd International Conference on Bioinformatics and BiomedicalEngineering. ICBBE 2009.

[8] G. Zhang, L.V. Berardi,(2007) “An investigation of neural networks in thyroid function diagnosis,” Health Care Management Science,1998, (pp. 29-37.)

[9] V. Vapnik,(2012).Estimation of Dependences Based on Empirical Data, Springer, New York.

[10] Obermeyer Z,(2016). Emanuel EJ. Predicting the future— big data, machine learning, and clinical medicine. N Engl ; (375:12161219).

[11]BreimanL.(2001) StatisticalModeling: the two cultures.Stat Sci. ;16:199-231..

[12] Ehrenstein V, Nielsen H, Pedersen AB, Johnsen SP, Pedersen L. (2017) Clinical epidemiology in the era of big data: new opportunities, familiar challenges. Clin Epidemiol. ; 9:245-250

[13] S. Godara and R. Singh,(2016) "Evaluation of Predictive Machine Learning Techniques as Expert Systems in Medical Diagnosis", Indian Journal of Science and Technology, (Vol. 910).

[14] Sunila, Rishipal Singh and Sanjeev Kumar.(2016) "A Novel Weighted Class based Clustering for Medical Diagnostic Interface." Indian Journal of Science and Technology (Vol 9).

AUTHOR PROFILES

Ms.M.ANITHA completed her Master of Computer Applications and Masters of Technology. Currently working as an Assistant professor in the Department of Masters of Computer Applications in the SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. Her area of interest includes Machine Learning with Python and DBMS.

Mr. EJJIVARAPU NAGARAJU completed his Masters of Computer Applications. He has published A Paper Published on ICT Tools for Hybrid Inquisitive Experiential Learning in Online Teaching-a case study- Journal of Engineering Education Transformations, Month 2021, ISSN 2349-2473, eISSN 2394-1707. Currently working has an Assistant professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR (DT). His areas of interest include Artificial Intelligence and Machine Learning.

Ms. GUNAGANTI. GOWTHAMI is an MCA student in the Masters of Computer Applications at SRK Institute of

Technology, Enikepadu, Vijayawada, NTR District. She has a Completed Degree in B.Sc.(computers) from Sri Durga Malleswara Siddhartha Mahila Kalasala women's college in Vijayawada. Her areas of interest are DBMS, Java Script, and Machine Learning with Python.