

Thyroid data analysis using Proximity Relative Compressed Algorithm for Health Care Applications

¹Selvi.K, ²Kaladevi.A.C

¹Professor, Pawai College of Technology, Namakkal

²Professor/CSE, Sona College of Technology, Salem-5

³AP/CSE, Sona College of Technology, Salem-5

¹selvimidu@gmail.com, ²kaladeviac@sonatech.ac.in, ³ramyaperumal@sonatech.ac.in

ABSTRACT

Data Analysis is popular in various fields of health data. This study's purpose is to analyze the characteristics of the medical system, applications, analytical methods, and medical information. Machine learning is one of the main processing methods in all walks of our life. In this article, investigation scholars wish to work in thyroid disease prediction through a reference source. Two most popular thyroid diseases are classified in the current research of thyroid dysfunction i.e. hyperthyroidism and hypothyroidism. The author analyzes the naive Bayes models, decision trees, and perceptron's and radial basis function networks to compare the four categories. The above models' classification has produced a remarkable precision of the results. Compared to other models, decision tree is the best classification model. With this excellent classification of recycled logs and adjacent data, the Proximity Relative Compressed Algorithm (PRCA) implements support using Python-based organization.

Keywords: Health Care Application, supervised learning algorithms, kNN, PRCA, Data Set, Medical Data.

1. INTRODUCTION

One out of ten Indians suffer from a thyroid problem. The disease mainly affects women aged between 17-54 years. In the acute phase of thyroid attacks, hypertension, high cholesterol levels, heart problems, infertility and depression are bound to manifest. Electronic Health Records (EHRs) contain information about individuals including observations, laboratory tests, diagnostic reports, medications, procedures, digital storage of patient identification information

and allergic information. The thyroid hormone is produced by the thyroid gland which is one of the endocrine glands.

The main function of this hormone is to take in calories, to promote metabolism in the human body, while at the same time reducing the excess secretion of other hormone glands. Thyroid identification problem from thyroid disease test report and hard work can only diagnose it through experience and knowledge. Traditionally, there are two methods, a blood test by a laboratory technician and other symptoms, its symptoms, the doctor's diagnosis and thyroid examination based on the patient's prognosis. For large data sets because it is not easy for everyone to check the results from the report forecast. Thyroid diseases are the main task is to predict with accuracy better.

Artificial intelligence can summarize sequential data from a data set. There are a lot of applications that use machine learning. Machine learning methods are used in key applications such as claims, hygiene, malicious identification, capital formation, political decision making, the past knowledge, the judgment. Therefore, it is necessary to explore computer vision.

Thyroid hormone secreted by the thyroid gland regulates the body's metabolic rate. Breakdown of thyroid hormones can lead to thyroid disease. The thyroid is a disease caused by the malfunction of endocrine gland. Main hormones released by endocrine glands in the blood are thyroxin (T4) and triiodothyroxine (T3).

Hyperthyroidism (thyroid function) refers to the condition where excess thyroxine is released by the thyroid gland. In this case, the body's metabolic rate increases greatly which results into a sudden loss of weight, fast or irregular heartbeat. Thyroid hormone activity is there to regulate metabolic rate and affect growth. Thyroid disease or thyroid disease usually has two main problems. They have an overactive thyroid gland found in the blood — while the thyroid is dysfunctional and releases very little thyroid hormone into the blood — an excess of thyroid hormone in hyperthyroidism. If the existing work design is not balanced, then the LDA is done using the LDA algorithm, that is a major drawback (i.e. the number height) of different types of objects.. LDA overlap and LDA sample validation are the least problematic options.

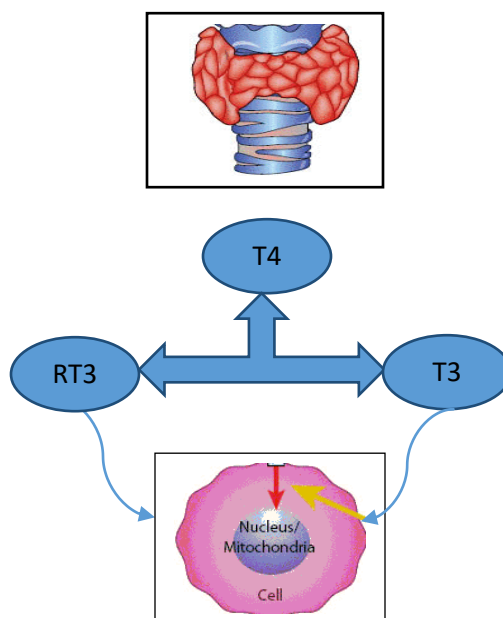


Figure 1 General Diagram for Factors that Affect Thyroid Function

Figure 1 explains the article is divided into three main parts, as followed by T4, T3, RT3, a brief introduction to the practice of the column, at the relationship between different data predictions based on the highlight. Section 3 Thyroid diseases that are proposed through identification and data mining, then collected information will include submitted to research. At the end of this chapter, the classification used in the calculation method and the evaluation of scenarios are discussed. Finally, the research work and its inherent problems and the writer put forward the investigation results to conclude.

2. RELATED WORK

Previous evidences regarding the effect of radiation [1] on thyroid immunity is under discussion. Non-Thyroiditis and Non-Obesity Diabetes (NOD) are Vulnerable to Iodine Treatment for 8 Weeks-H2 (H4) in mice, it has recently been documented that autoimmune thyroiditis increases with a low dose of whole-body radiation therapy (0.5 grays).

The security issue is the major issue in e-health care Telecommunications nowadays. The intent of ICT provides the best solution for efficient communication. They mainly focused on data privacy, CDMA encryption and decryption with less security [2]. They mainly focused on health care quality information and laws determining the medical institution. This review method medical health care information system supports Medical Service of the Electric Company of Caracas (MSECC), but patient data access and health care information low quality issued [3].

This method solves some of the challenges in this paper based on social psychiatric services as an imposed tool [4]. The medical health care services mainly focused many long-term health care homes need to improve the workload, reduce costs and improve the quality of medical services in some health resort, software technology assistance began to improved [5]. This paper focuses on monitoring a proposed work based on public primary care health data and medical health care Health covers the full cycle, but some health issues are found in central data management issued [6].

A health monitoring system based on Android apps focuses mainly on a different area and. This paperwork is a powerful tool with processing capability that supports storing and visualizing biomedical signals in the Android operating system. But recent issues centered on some are misusing the data view to another person [7].

Health institutions [8] Foundation is to raise awareness of the first stage of development. System-saving medical technology provides the necessary approaches to the management of files. Neural Radial Basis Function (RBF) [9] system is used to identify the categories of thyroid problems. Many modes of the human body are involved in protein production in the human body in response to the regulation of the human body [10]. The semi-automatic segmentation method of classification and analysis of the thyroid gland [11] are documented in the database. Fine Needle Aspiration (FNA)[12] identifies thyroid cancers in at least 0-45 compared to the level of cancer incidence. It focuses on high quality, medium quality and other areas. The effects of low levels of radiation exposure on the thyroid cancer system remain unclear [13].

Biopsy in clinical practice is required to be done to identify benign thyroid nodules. Therefore, the need to provide clinically relevant diagnostic information arises. This increase in a filler and non-invasive imaging tool to reduce the rate of unwanted biopsies of the thyroid gland has not been eliminated unnecessarily more than significantly more harmful nodules [14]. Thyroid ultrasound is a widely used medical technique for diagnosing thyroid tuberculosis. However, the lower the contrast, the noisier and the different appearance of skin nodules, the more difficult it is to diagnose and detect skin nodules. The structure of the Convolutional Neural Network(CNN) as a multi-functional waterfall uses ecological information on thyroid skin nodules. The structure is built with accurate and detailed ground true labels with many clinically diagnosed thyroid ultrasound images. Average Accuracy Malignant and benign thyroid

tuberculosis Diagnosis rate 5% Average lower-level error is better than normal cell neural network.

3. Materials and Methods

The therapeutic system and serum levels of free thyroxine (FT4) and Thyroid-Stimulating Hormone (TSH) are shown with the relevant data (Figure 2).

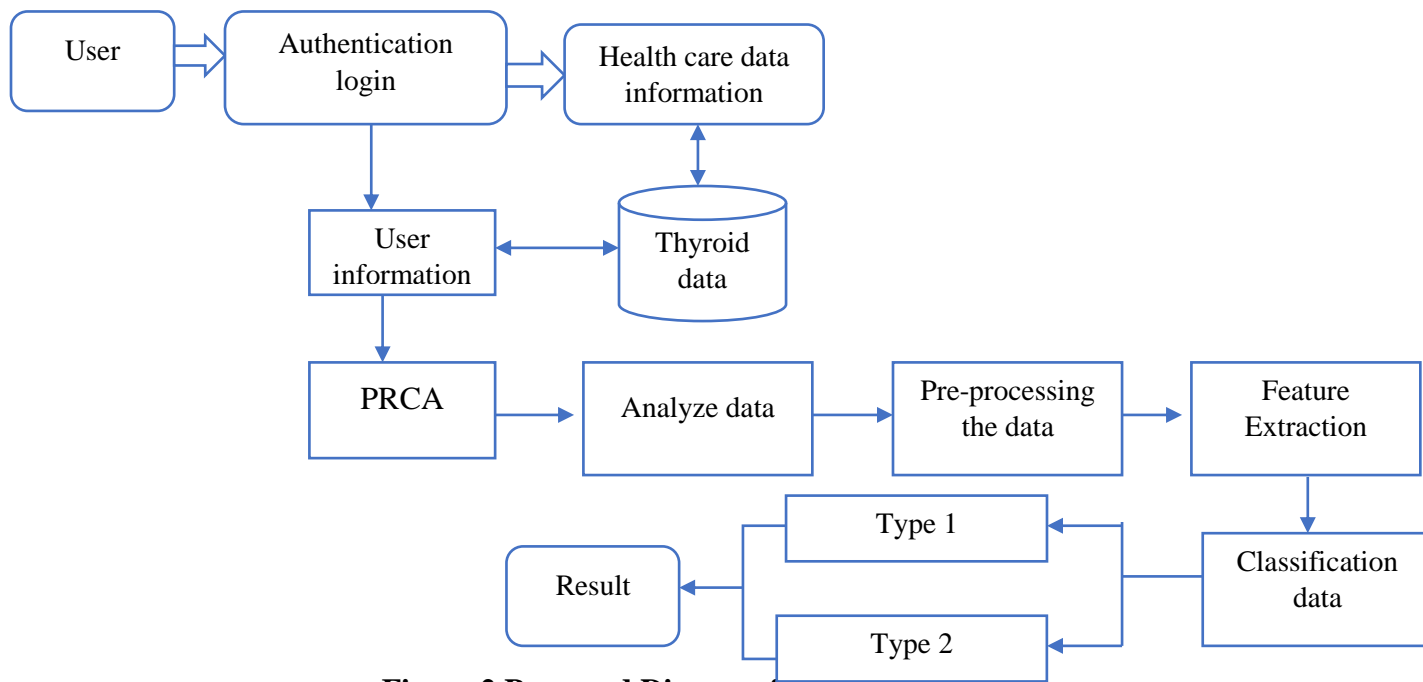


Figure 2 Proposed Diagram for Health Care System

The users can search for thyroid data files before entering the data server and import health information. This allows users to keep their medications up to date and use the protected email or data transfer method without the need for physical achievement of thyroid test results.

Thyroid treatment details are included in Figure 2 explains the and whether treatment time corresponds to serum FT4 and DSH values corresponding to those levels. In contrast, thyroid users in different countries deviate from the well-known equation of ng / dL and subsidiaries and have the opportunity to adjust the method of using picomols / L of serum T4 measurement using different measurement pattern.

3.1 Authentication Users

Direct recording of user activity and different settings will be part of the user roles. This processor system aims to protect the server configuration and easily test various users and functions using individual users' attribute files. With thyroid data, the user can view all the

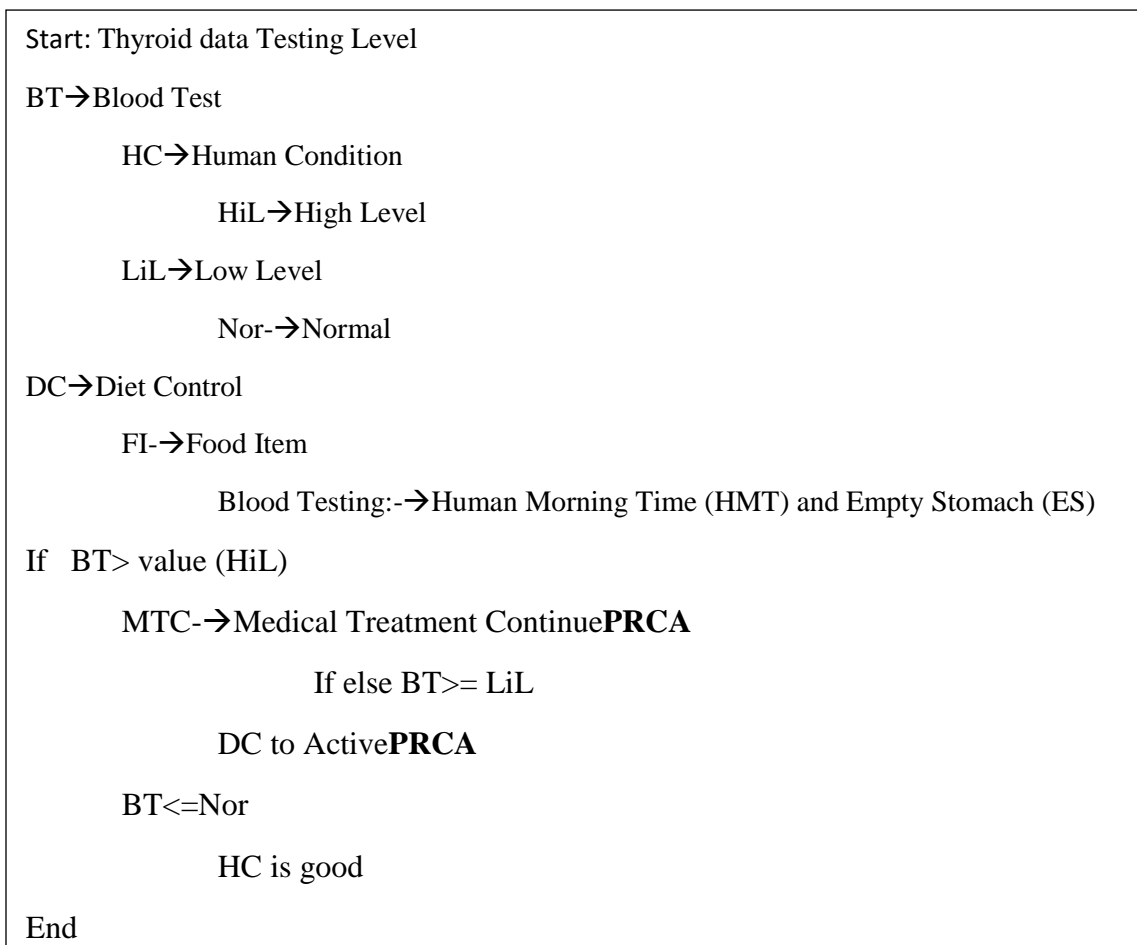
details and register the medical monitoring system to enter all the thyroid information in the registration form.

3.2 Health Care Data Information based on Thyroid data

The limit on the amount of thyroid data collected on display to display, computer-generated medical health information service providers. Many thyroid clinics monitor medical monitoring and clinical evaluations, and medical professionals are constantly monitoring the security of thyroid data in the database. There are many datasets in this work; the “new thyroid” data collection consists of two attributes and 215 events. Its serum is called total thyroxin (D4) and total serum thyroxin (D3); they also exclude the fact that there are two relevant attributes.

3.3 Proximity Relative Compressed Algorithm -PRCA

The proposed method of Proximity Relative Compressed Algorithm has a thyroid test level of data analysis to explain details below, and system support.



3.4 Preprocessing

The preprocessing of data is a significant step. It completely converts data into a readable format. The original data contains discrepancies, noise and loss of value. The data mining process and performance must begin to improve. For the thyroid disease data set, the missed value is replaced by the mean value.

$$\text{Mean}(x) = \text{sum of all values} \sum x / n \quad (1)$$

$$T(3) = \text{totalvalues}/\text{Avg} \quad (2)$$

3.5 Feature Extraction

Facility selection is also called machine learning attribute selection. It is the selection of related functions to understand the user technique that simplifies good templates. This exercise is used to reduce the amount of time spent. This feature is not the same as extraction. Characteristic assessment method is used to reduce the evolution of this thyroid database. Finally, the properties used to calculate the relationship between each standard are used to select the user file.

3.6 Classification data

Machine learning classification is a task, and the prediction is based on a set of parameter classes. In data mining, a classification is based on supervised learning algorithm. The main objective is to find the root of the problem in the classification of new data. Find groups that used different classification methods. The PRCA algorithm is used to improve the accuracy of the thyroid results in a data set.

4. Result and discussion

This tool connects with the design python tool anaconda which is simulated using a program created in the python language. The various factors and values are considered in simulation. In these parameters, the Anaconda web mining tools create a dataset that describes information.

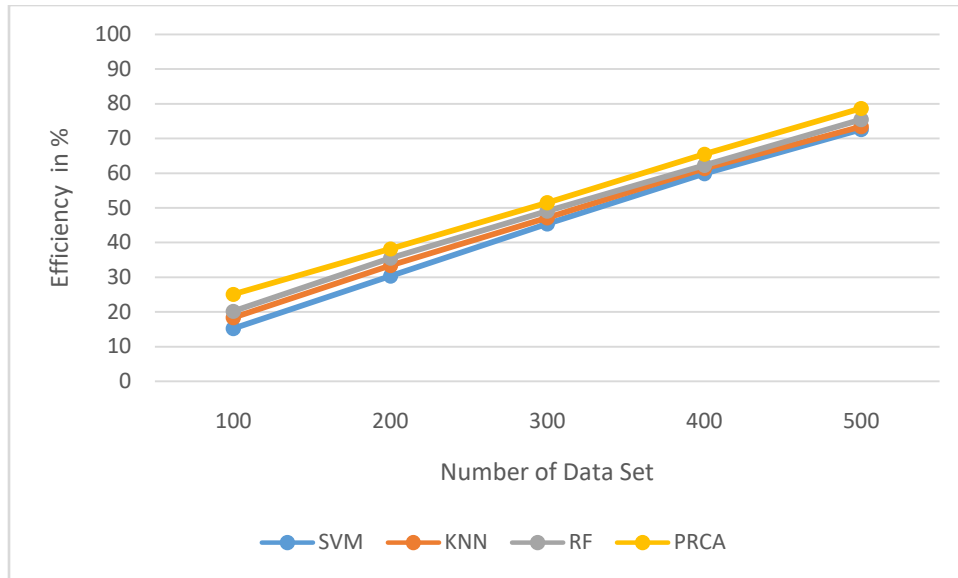


Figure 3 PRCA algorithm classification Performance

Figure 3 describes thyroid performance in health care using the Proximity Relative Compressed Algorithm (PRCA). The suggested algorithm gives the superior performance with 78.7 % compared to existing Support Vector Machine (SVM) algorithms with 72.6%, K-Nearest Neighbor (KNN) with 73.4%. With Random Forest (RF) with 75.4%.



Figure 4 Sensitivity performance

Figure 4 describes concerns about the sensitivity of the health care in the thyroid. People suffering from thyroid can cause excessive heat and sweating leading to sensitivity. Thus the

proposed Proximity Relative Compressed Algorithm (PRCA), achieves the extraordinary result of 68.6% sensitivity compared to existing algorithms Support Vector Machine (SVM) achieves 40.1% sensitivity, K-Nearest Neighbor (KNN) achieves 45.6% sensitivity and Random Forest achieves (RF) 55.5% sensitivity.

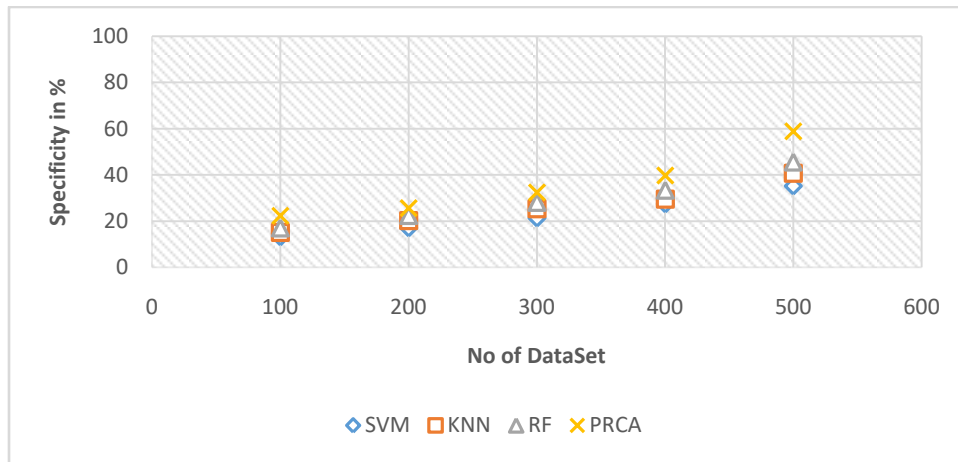


Figure 5 PRCA algorithm in Specificity

Figure 5 explains the thyroid specificity details. It is often considered a reference for assessing thyroid function. It is difficult to screen accuracy in primary care patients. Thus, the proposed Proximity Relative Compressed Algorithm (PRCA) achieves a high result of 58.9% compared to existing algorithms like Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF).

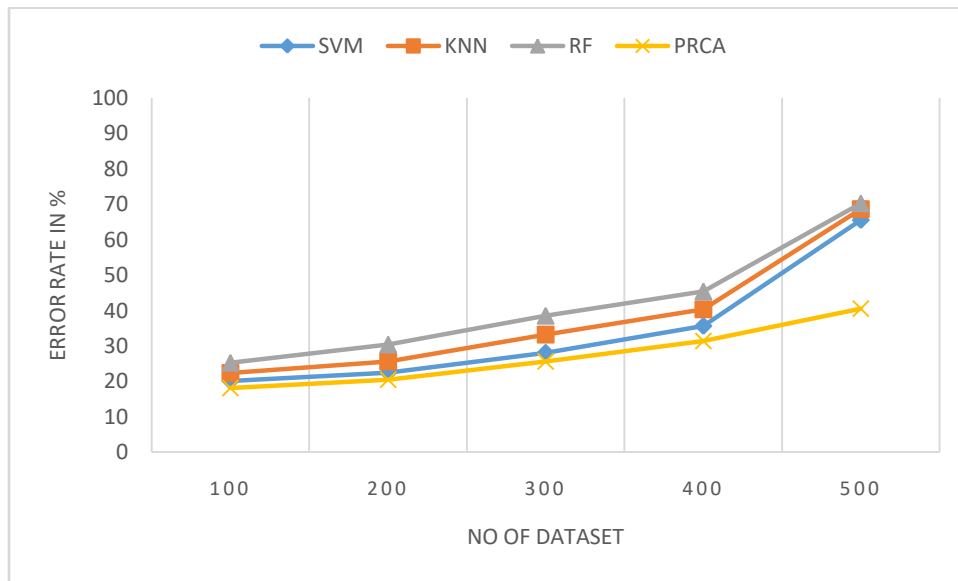


Figure 6 Health Care Data Prediction Error Rate (PRCA)

Figure 6 explains the error rate in the health care data prediction. Proximity Relative Compressed Algorithm (PRCA) gives a low error rate compared to existing algorithms Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Random Forest (RF). The suggested algorithm achieves a 40.5% of low error rate, and Support Vector Machine (SVM) gains 65.5% error rate, K-Nearest Neighbor (KNN) with 68.6% of the result, and finally, with Random Forest (RF) achieves 70.2% of high error rate. So, compared to the existing algorithm, suggested algorithm gives a low error rate.

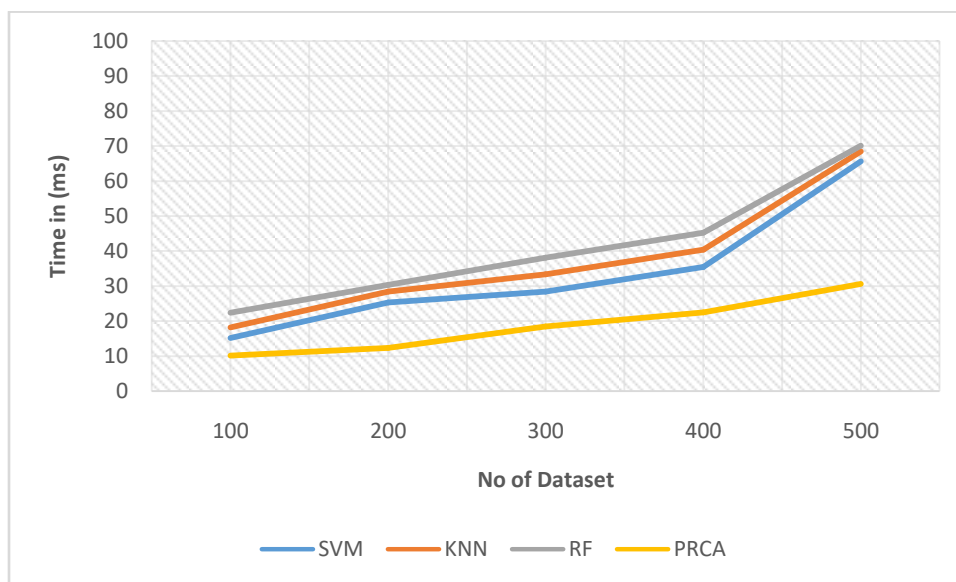


Figure 7 Performance of Time complexity

Figure 7 explains the time performance on Proximity Relative Compressed Algorithm (PRCA). This suggested algorithm achieves 30.6 ms of less-time performance than existing Support Vector Machine (SVM) algorithms with 65.6 ms, K-Nearest Neighbor (KNN) with 68.4 ms. Random Forest (RF). Finally, the proposed PRCA method proved to be the best.

5. Conclusion

The machine learning and data mining have been developed to analyze various sources of medical data sets. Increased accuracy and reliability of the diagnosis of the disease is a fundamental aspect of health data collection. Proximity Relative Compressed Algorithm (PRCA) is used to classify the thyroid diseases and their subtypes efficiently. Additional features are used to identify subtypes of thyroid RT3 and Basel Metabolic temperature. Thyroid result performance

in health care is augmented by using the Proximity Relative Compressed Algorithm (PRCA). Because the suggested algorithm gives the superior performance with 78.7 % compared to existing Support Vector Machine (SVM) algorithms with 72.6%, K-Nearest Neighbor (KNN) with 73.4% and Random Forest (RF) with 75.4% respectively.

Reference

1. Y. Nagayama, T. Ichikawa, O. Saitoh and N. Abiru, "Induction of Late-onset Spontaneous Autoimmune Thyroiditis by a Single Low-dose Irradiation in Thyroiditis-prone Non-obese Diabetic-H2h4 Mice," in *Journal of Radiation Research*, vol. 50, no. 6, pp. 573-577, Nov. 2009, doi: 10.1269/jrr.09067.
2. D. S. Venkateswarlu, K. S. Verma and K. S. R. A. Murthy, "e-Health networking to cater to Rural Health Care and Health Care for the Aged," 2007 9th International Conference on e-Health Networking, Application and Services, Taipei, Taiwan, 2007, pp. 273-276, doi: 10.1109/HEALTH.2007.381649.
3. S. Armengol, R. Mijares-Seminario, R. Silva-Bustillos, A. Monzón and J. Alfonzo, "Monitoring and evaluation of an occupational health program from "Electricidad de Caracas" and its Subsidiaries. Caracas-Venezuela," 2011 Pan American Health Care Exchanges, Rio de Janeiro, Brazil, 2011, pp. 446-448, doi: 10.1109/PAHCE.2011.5871947.
4. S. Manabe, N. Ashida, T. Takemura, K. Murase and N. Nishiura, "ICF-based community mental health care," *Proceedings. 6th International Workshop on Enterprise Networking and Computing in Healthcare Industry - Healthcom 2004 (IEEE Cat. No.04EX842)*, Odawara, Japan, 2004, pp. 110-113, doi: 10.1109/HEALTH.2004.1324487
5. P. Huang, C. Lin, Y. Wang and H. Hsieh, "Development of Health Care System Based on Wearable Devices," 2019 Prognostics and System Health Management Conference (PHM-Paris), Paris, France, 2019, pp. 249-252, doi: 10.1109/PHM-Paris.2019.00049.
6. S. J. A. Garcia, R. A. da Luz Santos, P. S. de Avelar, R. Zaniboni and R. Garcia, "Health care technology management applied to public primary care health," 2011 Pan American Health Care Exchanges, Rio de Janeiro, Brazil, 2011, pp. 250-253, doi: 10.1109/PAHCE.2011.5871898.
7. Z. Rebolledo-Nandi, A. Chávez-Olivera, R. E. Cuevas-Valencia, A. Alarcón-Paredes and G. A. Alonso, "Design of a versatile low cost mobile health care monitoring system using

- an android application," 2015 Pan American Health Care Exchanges (PAHCE), Santiago, Chile, 2015, pp. 1-4, doi: 10.1109/PAHCE.2015.7173334.
8. J. A. Villanueva and F. M. Martinez, "Toward a Health Care Technology Management Knowledge Base," 2009 Pan American Health Care Exchanges, Mexico City, Mexico, 2009, pp. 127-129, doi: 10.1109/PAHCE.2009.5158381.
 9. C. Chang, Y. Lei, C. Tseng and S. Shih, "Thyroid Segmentation and Volume Estimation in Ultrasound Images," in IEEE Transactions Biomedical Engineering, vol. 57, no. 6, pp. 1348-1357, June 2010, doi: 10.1109/TBME.2010.2041003.
 10. M. A. Savelonas, D. K. Iakovidis, I. Legakis and D. Maroulis, "Active Contours Guided by Echogenicity and Texture for Delineation of Thyroid Nodules in Ultrasound Images," in IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 4, pp. 519-527, July 2009, doi: 10.1109/TITB.2008.2007192.
 11. E. N. K. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger and T. Kuwert, "Quantification of Thyroid Volume Using 3-D Ultrasound Imaging," in IEEE Transactions on Medical Imaging, vol. 27, no. 4, pp. 457-466, April 2008, doi: 10.1109/TMI.2007.907328.
 12. K. K. Delibasis, P. A. Asvestas, G. K. Matsopoulos, E. Zoulias and S. Tseleni-Balafouta, "Computer-Aided Diagnosis of Thyroid Malignancy Using an Artificial Immune System Classification Algorithm," in IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 5, pp. 680-686, Sept. 2009, doi: 10.1109/TITB.2008.926990.
 13. T. Ohira et al., "Absorbed radiation doses in the thyroid as estimated by UNSCEAR and subsequent risk of childhood thyroid cancer following the Great East Japan Earthquake," in Journal of Radiation Research, vol. 61, no. 1, pp. 243-248, Dec. 2019, doi: 10.1093/jrr/rrz104.
 14. M. Mehrmohammadi et al., "Comb-Push Ultrasound Shear Elastography (CUSE) for Evaluation of Thyroid Nodules: Preliminary In Vivo Results," in IEEE Transactions on Medical Imaging, vol. 34, no. 1, pp. 97-106, Jan. 2015, doi: 10.1109/TMI.2014.2346498
 15. W. Song et al., "Multitask Cascade Convolution Neural Networks for Automatic Thyroid Nodule Detection and Recognition," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1215-1224, May 2019, doi: 10.1109/JBHI.2018.2852718.