Randomized Vector Network Model for Thyroid Prediction Using Relief And Lasso Feature Selection Approaches

Maruthi Prasad*, Santhosh R.
Department of Computer Science and Engineering, Faculty of Engineering, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
Emails: karpagam.publication@gmail.com; santhoshrd@gmail.com

Abstract

The studies' primary aim is to help the research scholars as a source who would like to research in the thyroid disease detection region. UC Irvin knowledge discovery provides databases files for the machine learning archives' thyroid dataset. Here, a random vector network model (RVNM) is proposed to perform classification tasks. The proposed model integrates the prior dataset information regarding the samples to train the more effective classifier. This cascaded random vector network model helps in thyroid disease prediction. The evaluation process is performed to predict and determine the respective performance concerning accuracy. The intuition is provided in this research, like forecasting the thyroid disease; it also calls attention to the process of using a Randomized Vector Network Model (RVNM) as a medium for classification. The simulation is done in the MATLAB 2020a environment and establishes a better trade-off than various existing approaches. The model gives a prediction accuracy of 96.1% accuracy compared to other models and shows a better trade than others.

Keywords: Thyroid disease; classification; randomized vector; prediction accuracy; attention

1. Introduction

Out of ten, at least one person in India is affected by thyroid disease. This thyroid disorder problem affects women between the age group of 17 and 54. Moreover, cardiovascular complications, BP increases, the cholesterol level increases, and attains depression and minimized fertility [1] is the extreme phases of thyroid disease. There are two active thyroid hormones called T4 and T3 generated with the help of the thyroid gland to manage the rate of metabolism of the human body. These hormones are essential for the complete yielding of energy, regulation, and the generation of proteins in the arms of human body temperature for every cell to function and every tissue and organ in the correct way [2] and [3]. The functional behaviour of the thyroid disease represents the concept of diagnosing the thyroid disease and the therapy. It is an essential element in many thyroid disorders where the fundamental classifications of thyroid disorder are hypothyroidism, euthyroidism, and hyperthyroidism. It refers to thyroid hormones' defective levels, average and excessive stages. The hypothyroidism stage has occurred because of the absence of thyroid hormone production and the deficient alternate therapy [4]. The euthyroidism state refers to the generation of thyroid hormones commonly and the average levels at the cellular stages with the help of the thyroid gland. A hyperthyroidism stage is the indication of the clinic because of the extra circulation and the intracellular thyroid hormones [5].

Some health care practitioners have the general concern to cure the disease and diagnose the condition without error at the correct time, which is essential for the patient. Moreover, the available medical report can be created by some state of the art diagnosis techniques with extra information related to symptoms. There are various questions like "Due to the thyroid, which age group of people are damaged?", "What would be the suitable treatment for the thyroid disorder?", "What are the reasons for the thyroid to affect humans?" and so on. The answers need to find for these questions to implement the machine learning techniques. After processing and implementing health care data with some methods, these results can be used to diagnose and provide the treatment for the disease more effectively and correctly with good decision-making and lowers the risk of death [6]. The massive volume of data needs to be managed with the help of machine learning approaches. The classification models are opted to classify and distinguish the data classes. However, classification processes help handle

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

categorical and numerical values [7]. There are two steps in the classification model. In the first step, a model related to some training data is established. In the second step, an unspecified tuple is provided to the model for the classification to class label [8].

The classification has a significant impact on humans' life. The various classification approaches are compared, which is essential and contain good dependence on the properties of the datasets. Some algorithms like decision tree, k-nearest neighbour, and logistic regression attain the respected place in the statistics community for the classification issues [9]. A small work needs to be carried out in the classification techniques for the patients based on the literary works and research reviews that are pruned with the help of thyroid disease. The classification methods which are utilized are well-known techniques [10]. This research describes the need for classification machine learning algorithms to concentrate on the problems mentioned above in thyroid disease [11]. This research describes the construction, testing, and training of data [12] – [15]. Each method possesses various steps to explain, and the accuracies of different techniques are compared that are utilized in the prediction. This research concentrates on modelling a cascaded random vector network model (RVNM) for thyroid prediction by reducing the complexities encountered in data acquisition. The features from the thyroid dataset are fed as an input to the cascaded random vector model to generate the corresponding prediction values, and the feasibility of the model helps to construct a better prediction approach and to give the optimal diagnosis results.

The remainder of the work is structured as follows: Section 2 analyses existing approaches and issues; section 3 elaborates the proposed cascaded random vector network methodology. The numerical outcomes of the anticipated model are provided in section 4, followed by the research summary in section 5.

2. Related work

Various researchers handle thyroid disease identification in the literature and are grateful for using hormonal parameters and the patients' data like gender and age. Moreover, the classification of machine learning and the prediction models are used in some studies when other techniques utilize deep neural network techniques. Lin et al. [16] used the dataset UCI repository in the first group to classify thyroid disorder with the help of a decision tree model. Specifically, MLTDD is developed as a machine learning tool to diagnose thyroid disease, generating the intelligent prediction of thyroid gland disorder. The research provides the complete accuracy of 98.7% and 99.8% for testing. There are some machine learning approaches like MLR, SVM, DT, and NB are utilized by the authors in [17] and [18] to perform the comparison on the thyroid disorder diagnosis. Its outcomes provide precision that is the same as 99.23%. The decision tree provides better performance and needs to support the thyroid gland's successful detection.

Significantly, research is conducted on the 244 subjects that suffer from various pathologies to examine patients' thyroid state that considers few parameters of hormone and age. The authors predict thyroid disorder [19] and [20] by conducting research using various data mining approaches and identifying the correlation among T3, T4, and TSH features with hypothyroidism or hyperthyroidism. Moreover, the SVM, k-NN, ID3, NB are used as the data mining algorithms by the authors by using them to the UCI data archive of the public dataset. In particular, the neural networks such as PNN, MLP, FTDNN, CFNN, GRNN by the authors in the research [21] – [25] for the diagnosis types of thyroid disease. The study provides the results that the neural network provides more specific results and classifies the thyroid pathology concerning the hormonal parameters. The authors conducted the research to diagnose hypothyroidism and hyperthyroidism, considering the two most prevalent thyroid diseases [26]. There are two techniques used for the classification: neural networks and multinomial logistic regression models. There are 310 patients utilized in this research, and the models are considered the input demographics and the hormone parameters in this case. The outcomes provide better performance by the neural network approach in all patients with an accuracy of 96%, rather than multinomial LR with a mean value of about 91.4% accuracy. This research varies from the mentioned techniques. The first technique focuses on predicting thyroid disease treatment, which is the best of our knowledge [27]. Henceforth, the vital aim of this research is to utilize the complete data gathered in time on the patient for prediction if the LT4-based treatment requires maximizing or minimizing. Moreover, the various machine learning techniques are used and compared in the study in predicting the need of caring for patients affected by a thyroid disorder [28] – [30]. However, the research suggests a new group of features found based on the experience of endocrinologists in the treatment of the patient. Lastly, using a dataset to extract the actual data is the crucial contribution of this research. Integrating two sub-datasets helps build this dataset with information regarding the patient's medical history and current state.

3. Methodology

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

The goal of this proposed technique is to forecast the treatment concerning the thyroid disorder based on the historical and current data of the patient. The construction of the dataset is explained in this section primarily, and then the suggested features models are described later. Lastly, the machine learning techniques concentrate on performing the research and the relevant validation.

a. Data collection

From the "AOU Federico II" Naples hospital [5], the datasets need to be taken from patients related to thyroid disease to be treated for conducting the research. Moreover, the data sources need to be gathered from every patient like personal data like date of birth, age, pathology, sex, education, marital status, and profession; also details of family history, the physical features like weight, height, data concerned to possible pregnancies and mensuration for women, body mass index, and few considering alvus, diuresis, appetite are considered. In addition, some clinical data also needs to be considered like neck, skin, thorax, heart, extremities, eyes, and abdomen. The obtained datasets are deemed to incorporate two data sets concerning 800 patients. Secondly, the data source is obtained from the diary of doctors' visits. It provides all the related data of every patient regarding clinical visits and tests are conducted while visiting the doctor. The two data sources provide the data, and then it is combined into a single massive dataset with the help of a patients' ID used as an element. Consecutively, the cleaning work needs to be carried out. The management of all the uncorrected data and missing values is primarily performed. However, patients who have a single visit are eliminated from the dataset by looking at every patient's clinical history. They are not required for the evolution of the disease in this research.

Further, this proposed approach selects only the patients who suffered from hypothyroidism due to the examined medicine in predicting the disease. There are three macro-groups of pathologies in the gathered datasets. They are (a) Congenital hypothyroidism, (b) Hashimoto's thyroiditis with hypothyroidism, (c) hypothyroidism. The finally collected dataset has 247 patient data, like 51 men and 195 women and mean age of 46 years. In some circumstances, patients with hypothyroidism visit the hospital for more than a year or many appointments in a single year. In particular, a group of collected datasets has 2784 instances that refer to a specific patient and a particular visit while the essential clinical and physical data are deposited (See Table 1).

Table 1: Dataset attributes

Feature	Description	Type
Height	Patients' height (cm)	I
Body mass index	Patients' BMI	F
Age during visit	Patients' age during hospital visit	I
Pathology	Patients' thyroid disease	S
Pathologies severity	thyroid disease severity	S
Pathologies Causes	thyroid disease causes	S
Weight	Patients' weight (kg)	I
Gender	Patients' gender (Male-Female)	S
Familiar anamnesis: Thyroid Diseases	Thyroid disorder patients' family (yes-no)	B
Familiar anamnesis: Diabetes	Diabetics patients' family	B
Menarche	Patients age during menarche (female patients alone)	I
Menstruation	Woman period (female patients) (regular - irregular)	S
Pregnancies	Pregnancies count, (female patients)	I
interruptions in pregnancy	Interruption count during pregnancy (female patients)	I

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

Menopausal age	patients' age during menopause (female patients)	I
Appetite	patients' appetite rate (poor – good – regular – excessive – great - variable)	S
Bowel function	patients' bowel function (regular – irregular – constipation – frequent - variable)	S
Diuresis	patients' diuresis degree (regular – irregular – frequent – poor - variable)	S
TFH	TSH (Thyroid Simulating Hormone) is a pituitary hormone restores the thyroid gland to generate (T4) and (T3) converts the body's metabolism	F
FT3	FT3 thyroid hormone with iodine	F
FT4	FT4 Thyroxine is a thyroid hormone that is partly made up of iodine	F
Thyroglobulin	Thyroid cancer therapy efficiency.	F
Ab<>	Thyroid anti-bodies are the immune system elements misdirected against the thyroid gland or some fundamental reasons for its general functions.	F
AbTg	Anti-body thyroglobulin	F
AbTPO	Anti- thyroperoxidase anti-bodies	F
LT4	Levothyroxine specifies L-thyroxine using hybrid hormone thyroxine (T4). Used for heal and avoid some thyroid cancers types (doses/ week)	I
LT4 treatment	LT4 treatment trend (higher – lower – stable - others)	S

S- string, I- Integer, F-Float, B-Boolean

b. Data acquisition

A considerable volume of gathered data is collected through surveys, experiments, the internet, etc. Moreover, the noise, missing values, and distortions data are often used. The integrated dataset also has the null or missing values used in this research. Some common approaches like deletion and imputation are utilized to handle the missing values. Table 1 provides the range of importance in the dataset. The data needs to be standardized or normalized before the machine learning algorithms are used. This issue is solved with the help of the imputation approach. The conversion of information to the mean of (μ) and a standard deviation (P) of 1 is done by standardization. The conversion formula of (8) is provided as follows in Eq. (1):

$$\text{Standardization}, X = (X - \mu)/\sigma \quad (1)$$

c. Feature selection

The machine learning approach considers the technique of feature selection as essential and the better features for the process of classification that requires extract. This process reduces the execution time. There are two algorithms selected in this proposed approach. They are (i) Relief feature selection and (ii) Least Absolute Shrinkage and Selection Operator.

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

i) Relief Feature selection

Relief is the algorithm related to the selection attribute that provides the weight to the entire attributes in the datasets. The weight needs to be altered slowly [64]. The essential features should have a high, and the available features should have fewer weights is the goal of this approach. Relief utilizes the same methods like in K-nearest neighbour to evaluate the weights of the feature. Kira and Rendell [65] designed a well-defined algorithm of feature selection techniques. R_i is suddenly used for the selected instance. There are two nearest neighbours in the Relief technique, where the first one uses a similar class to the closest hit, H. The second uses the opposite class referred to nearest miss M. The consistencies need to be computed W[A] adjusted based on the feature′s R_i, H and M values. The performance value W[A] is lowered when a massive differentiation between R_i and H occurs, which is undesirable. Contrarily, feature A needs to be utilized to differentiate the various classes. The weight W[A] is improved when a vast differentiation between R_i and M for feature A is available. These processes are pursued four times when parameter m needs to adjust (Table 2).

Table 2: Selected features with Relief model

Feature	Code	Score
Body Mass Index	Patients BMI	0.20
Weight	Weight in kg	0.85
Pregnancies	Count of pregnancies	0.0233
TFH	T4 and T3 TSH	0.582
FT3	Thyroid hormone	0.544
FT4	Thyroid hormone	0.0090
LT4	Levothyroxine	0.582

Table 3: Selected features with LASSO

Feature	Code	Score
Pathologies severity	Thyroid disease severity	0.0013
Menarche	Age during menarche	0.0806
Pregnancies	Count of pregnancies	0.0019
TFH	T4 and T3 TSH	0.0000
FT3	Thyroid hormone	0.0014
FT4	Thyroid hormone	0.0230
LT4	Levothyroxine	0.0317
AbTg		0.0115

ii) Least absolute shrinkage and selection operator algorithm

The shrinkage functionality and the operators' minimum selection are concerned with modifying the absolute value of the functions' coefficients. Few features have zero coefficient values, and some attributes are eliminated, having a time coefficient from the features' subsets. LASSO provides better performance for the feature value having small coefficients. The attributes having significant coefficients values are there to choose the features' subsets. LASSO has unwanted features. However, the above procedure is repeated many times to achieve the reliability of the feature that needs to be improved eventually, considering the more often identified attributes in the research. It is known as the randomized LASSO feature are initiated (Table 3). It needs to be established on a robust computer that utilizes parallel programming. The realization for the latest application is found where $q - i$ denotes the vector of concerned i^{th} sub-region keys.

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

d. Random Vector Network Model (RVNM)

Fig 1 depicts the flow diagram of the anticipated RVNM algorithm, and the below-given steps are performed using the model:

1) The extracted features from the input are provided to the initial layer select vector values during the training process.

2) The originally extracted features are provided input to the successive layers with the select vector values produced by the classifier model over the current layer. Thus, the privileged information is self-generated devoid of any added modalities.

3) The absolute values generated from the second layer (as provided in step 2) are then provided to form the third layer. The predictive outcomes from the successive layer classifiers are set as the privileged input for the next (third) layer. This process is repeated to construct a broader classifier model using the cascaded form. The parameters of every layer are fine-tuned and learned with the new information and corresponding values. This successive way for parameter-tuning leads the classifier of every layer (last) with superior parameters attain higher classification outcomes (Fig 2).

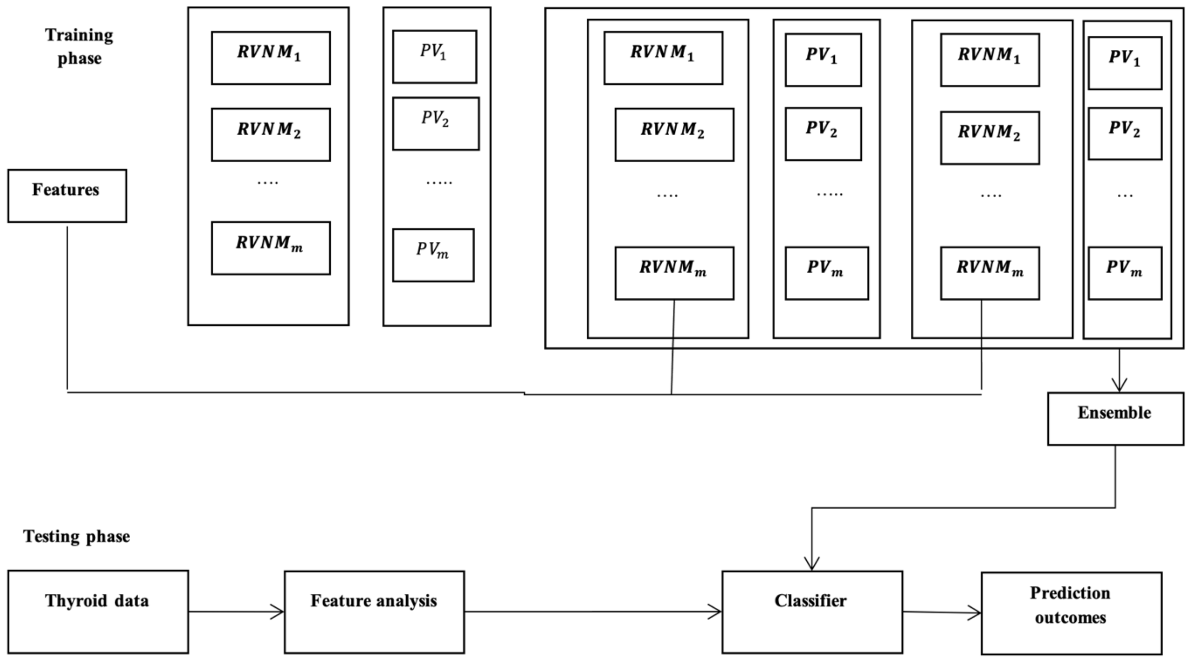

Figure 1: Random Vector Network Model architecture

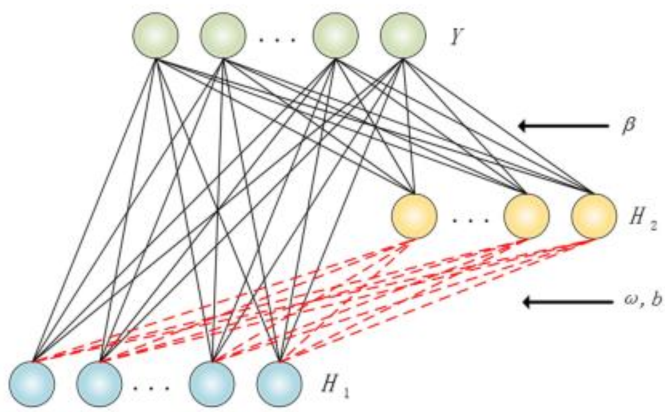

Figure 2: Internal architecture

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

4) The concept of the ensembling process is applied lastly to the final column of the classifier (last layer) in the network for categorization purposes. Here, the final classified label is determined using the maximal class score from the mean of every prediction outcomes.

5) In case of training, the hauled out original features from the input are given to the classifier's final column directly. Finally, it gives the diagnosis outcomes with the ensemble learning process.

The proposed model is beneficial and establishes a direct link among the outputs and inputs to evaluate the randomized variant of the neural networks' functional link. The enhancement network nodes among the input/output layers for the weight vectors and bias are allocated randomly. The interconnection weights from the input/enhancement nodes over the output layers are optimized using L2-norm regularization. Consider the provided data $\{(x_i, y_i)|x_i \in R^d, y_i \in R^m, i = 1, \dots, n\}$, where x_i is measured as feature vector, and y_i is considered as the related labels. Then, the process defines the H_i as the input matrix, and H_2 specifies the enhancement matrix as depicted in Eq. (2) & Eq. (3):

$$H_1 = \begin{bmatrix} x_{11} & \dots & x_{1d_1} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nd1} \end{bmatrix} \quad (2)$$

$$H_2 = \begin{bmatrix} g(\omega_1 . x_1 + b_1) & \dots & g(\omega d_3 . x_1 + b_{d_3} \\ \vdots & \ddots & \vdots \\ g(\omega_1 . x_n + b_1 & \cdots & g(\omega_{d_3} . x_n + b_{d_3} \end{bmatrix} \quad (3)$$

The enhancement nodes are produced using the activation function $g(\omega_j . x_i + b_j)$ through the random weight ω_j and bias b_j $(j = 1, \dots, d_3)$. The output weight β needs to be evaluated using Eq. (4):

$$H\beta = Y \quad (4)$$

Here, Y specifies the label matrix, H is a concatenated matrix, provided in the H form $H = [H_1, H_2]$. Here, β is evaluated using the L2 norm regularized using Eq. (5):

$$\beta = \left(H^T H + \frac{I}{C}\right)^{-1} H^T Y \quad (5)$$

The successive layers bridge the gap among the NN and the privileged information to integrate that information into the random vectors, and these random vectors are expressed as in Eq. (6) & Eq. (7):

$$\min \frac{1}{2} ||\beta||_2^2 + C \sum_{i=1}^{n} \mu_i \quad (6)$$

$$h(x_i)\beta = y_i - \mu_i, \ \forall\, 1 \le i \le n \quad (7)$$

Where μ_i is the training error vector, $h(x_i)$ specifies the concatenated vectors alike of H over the above equation. Some additional information is added to generate the newer training set $\{(x_i, p_i, t_i)|x_i \in R^{d_1}, P_i \in R^{d_2}, t_i \in R^m, i = 1, \dots, n\}$ where p_i refers to the information over the randomized vector. The generated vector needs to handle the objective function, which is alike of the existing SVM, and it is expressed as in Eq. (8) & Eq. (9):

$$\min_{\beta, \beta} \frac{1}{2} ||\beta||_2^2 + \frac{\in}{2} \left|\left|\bar{\beta}\right|\right|_2^2 + C \sum_{i=1}^{n} \mu_i \ (\tilde{\beta}, \tilde{h}\ (p_i)) \quad (8)$$

$$h(x_i)\beta = y_i - \mu_i\ (\tilde{\beta}, \tilde{h}\ (p_i); \forall 1 \le i \le n \quad (9)$$

Here, $\in$ it specifies the regularization coefficient, $\tilde{h}\ (p_i)$ is alike of $h(x_i)$ related to privileged information, and $\mu_i(\tilde{\beta}, \tilde{h}\ (p_i))$ specifies the proper functionality of the output weighted vector $\bar{\beta}$ over the feature space. It is expressed as in Eq. (10):

$$\min \frac{1}{2} ||\beta||_2^2 + \frac{\in}{2} \left|\left|\bar{\beta}\right|\right|_2^2 + C \sum_{i=1}^{n} \tilde{h}\ (p_i)\tilde{\beta} - \sum_{i=1}^{n} \lambda_i\ (h(x_i)\beta - y_i + \tilde{h}\ (p_i)\tilde{\beta}) \quad (10)$$

Here, λ_i specifies the lagrange multipliers $\lambda_i = [\lambda_1, \dots, \lambda_n]^T$. Moreover, using specific computational factors to measure the lagrangian's saddle points, the equation is formed as Eq. (11) to Eq. (13):

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

$$\beta = H^T\lambda \quad (11)$$

$$\tilde{\beta} = \frac{1}{\in}(\tilde{H}^T\lambda - \tilde{H}^T C1) \quad (12)$$

$$h(x_i)\beta - y_i + \tilde{h}(p_i)\tilde{\beta} = 0 \ \forall 1 \le i \le n \quad (13)$$

By evaluating the above Eq. (13) in Eq. (14), we get,

$$(HH^T + \frac{1}{\in}\tilde{H}\tilde{H}^T)\boldsymbol{\lambda} = Y - \frac{C_1}{\in}\tilde{H}\tilde{H}^T \quad (14)$$

With the integration of Eq. (13) and Eq. (14), the closed-form solution is attained in Eq. (15):

$$\beta = H^T(HH^T + \frac{1}{\in}\tilde{H}\tilde{H}^T)\ (Y - \frac{C_1}{\in}\tilde{H}\tilde{H}^T) \quad (15)$$

However, $\frac{I}{C}$ is added to the above Eq. (16), therefore to eliminate the singularity and to fulfil the random vector models' stability which provides the following solution:

$$\beta = H^T(HH^T + \frac{1}{\in}\tilde{H}\tilde{H}^T + \frac{I}{C})\ (Y - \frac{C_1}{\in}\tilde{H}\tilde{H}^T) \quad (16)$$

The output function is attained with the following Eq. (17):

$$f(x) = \ h(x)\beta = H^T(HH^T + \frac{1}{\in}\tilde{H}\tilde{H}^T + \frac{I}{C})\ (Y - \frac{C_1}{\in}\tilde{H}\tilde{H}^T) \quad (17)$$

The attained output function during the testing data is computed with Eq. (17) using testing data of training samples.

4. Numerical results

The evaluation of the efficiency and accuracy of the machine learning technique is performed with the help of performance indicators. When a person is categorized as HD, the positive classification is considered. The person comes under negative classification when the person is not classified as with HD.

TP = True Positive (If the technique is accurately found as HD).

TN = True Negative (If the technique recognized the opposite class like patients with no heart problems truly).

FP = False Positive (If the technique inaccurately recognized the HD patients that is finding the non-HD patients as HD patients)

FN = False Negative (If the technique inaccurately recognized the opposite class like HD patient as the average patient).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (19)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (20)$$

$$F1 - \text{score} = \frac{2 * (\text{precision} * \text{recall})}{(\text{Precision} + \text{Recall})} \quad (21)$$

$$\text{False Positive Rate} = \frac{FP}{FP + TN} \quad (22)$$

$$\text{False Negative Rate} = \frac{FN}{TP + FN} \quad (23)$$

$$\text{Negative Predictive value} = \frac{TN}{TN + FN} \quad (24)$$

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

The proposed classifiers' performance is better than the existing research in this research (Table 4). Moreover, some restrictions are dependent on the particular approach called feature selection. Illustrate with an example, highly accurate outcomes are produced when high reliance on Relief has occurred in this type of case. In addition, there is an adverse impact when a more level of missing values is available in the datasets. The process of addressing the problems via the appropriate methods and using other datasets are demonstrated when this model is utilized. When the missing values are a little remarkable, this model needs to manage the problems. Further, a large dataset needs to develop the model more specific even though the training dataset is considerably extended in the proposed approach (See Fig 3 to Fig 6).

Table 4: Prediction outcomes

Iterations	TP	FP	FN	Precision	Recall	Accuracy
1	590	8	19	98	97	95.9
2	576	5	31	99	95	95.2
3	630	20	51	97	92	90.2
4	600	35	18	94	97	95.1
5	565	45	10	92	98	96.5
6	630	25	10	99	98	98.2
7	600	30	21	96	96	95.2
8	600	75	12	95	98	97.7
9	625	60	13	89	98	97.9
10	555	13	32	91	95	97.2
11	600	25	56	97	92	97.1
12	580	35	41	96	93	97
Overall	4190	376	314	95.25	95.75	96.1

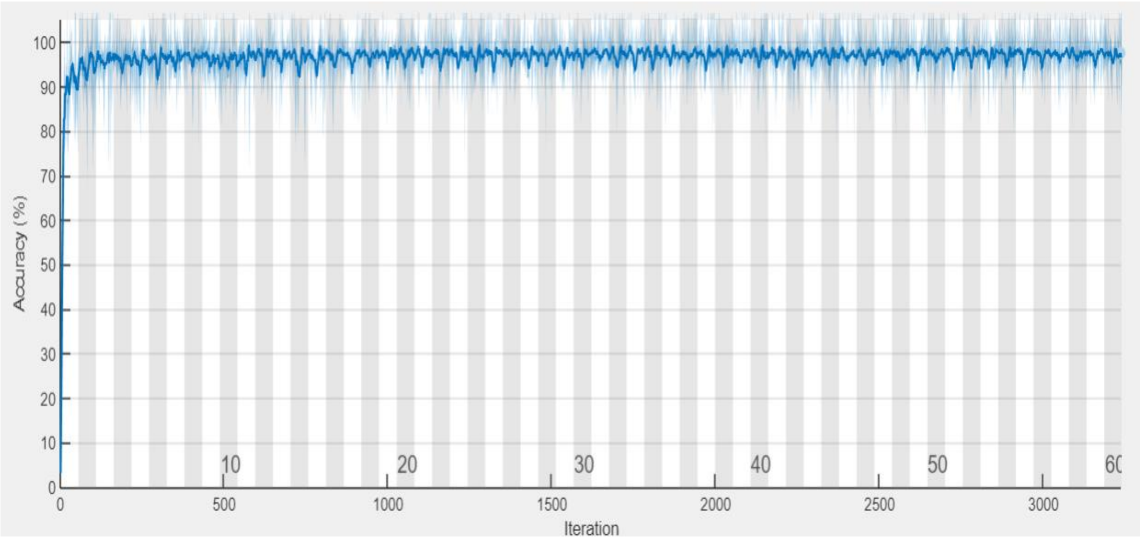

Figure 3: Accuracy computation

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

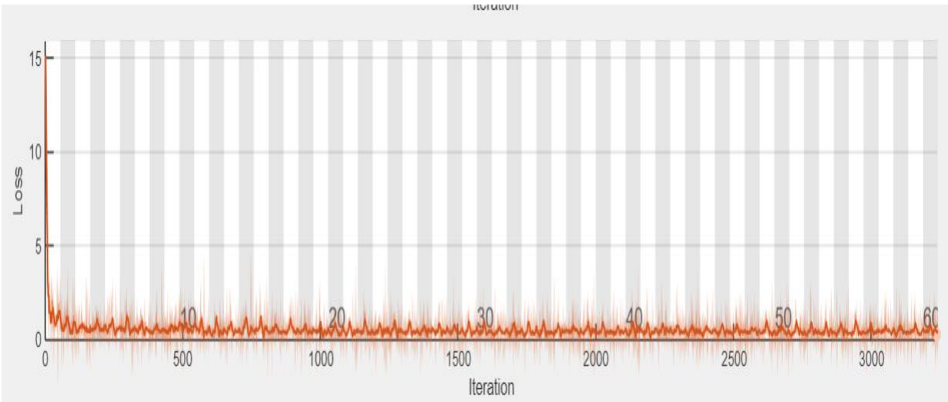

Figure 4: Loss computation

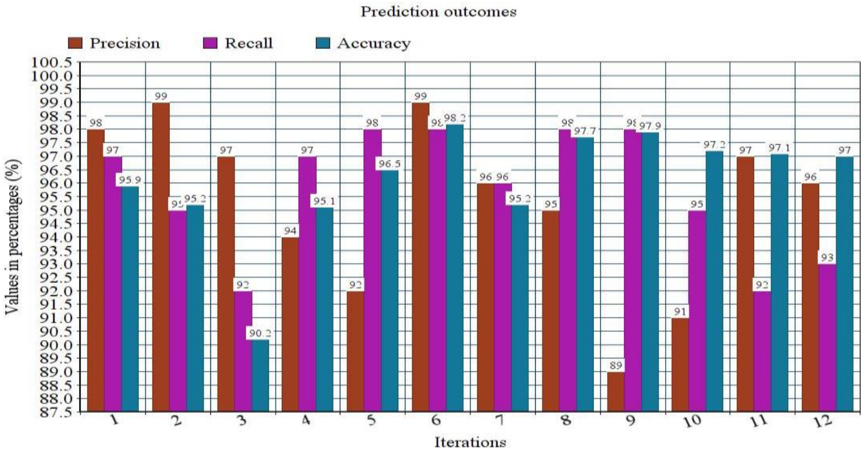

Figure 5 Precision, recall and accuracy comparison

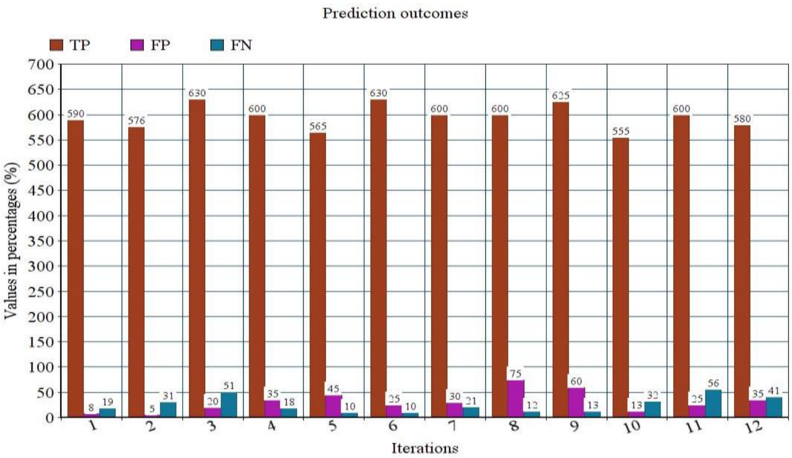

Figure 6: TP, FP and FN comparison

Table 4 compares various performance metrics like TP, FP, FN, precision, recall and accuracy. The proposed model gives average prediction outcomes of 95.25%, 95.75% recall and 96.1% accuracy. This work performances a successive iteration for 100 epochs where the samples of 12 iterations are provided. The average TP for 12 iterations are 4190, FP is 376, and FN is 314. The model performance is higher compared to various existing approaches. Based on these outcomes, it is observed that the proposed RVNM works well in thyroid prediction and establishes a better trade-off. Similarly, the prediction quality is improved by adopting superior feature

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

selection approaches like Relief and LASSO. The former model selects seven features while the latter selects eight features. Those selected features are fed as an input to the classifier, and better prediction outcomes are attained.

5. Conclusion

The models need to be developed, which helps predict the different stages of thyroid in the present situation with the help of machine learning. The available data and the generation of data are increased a chance day by day for the computer scientists to predict and analyze the data sets which creates the human life good and convenience. This research is concerned with the aim. The classification and prediction of the data are based on the data set itself, and the proposed RVNM is used along with the Relief and LASSO feature selection process. The proposed model gives an average prediction accuracy of 96.1% is substantially higher than the existing approaches. The proposed method does not consider the data set to predict the thyroid. It needs to be regarded as the future scope and calculate the accuracy with the help of other network models. Henceforth, the obtained accuracy is satisfactory according to the data set used in this approach.

References

[1] Kondo, K. Takagi, M. Nishida, T. Iwai, Y. Kudo, K. Ogawa et al., “Computer-aided diagnosis of focal liver lesions using contrast-enhanced ultrasonography with perflubutane micro-bubbles,” IEEE Transactions on Medical Imaging, vol. 36, no. 7, pp. 1427–1437, 2017

[2] Feng, F. Yang, X. Zhou, Y. Guo, F. Tang, F. Ren et al., “A deep learning approach for targeted contrast-enhanced ultrasound based prostate cancer detection,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 16, no. 6, pp. 1794–1801, 2018.

[3] Rizzo, M. Tonietto, M. Castellaro, B. Raffeiner, A. Coran, U. Fiocco et al., “Bayesian quantification of contrast-enhanced ultrasound images with adaptive inclusion of an irreversible component,” IEEE Transac1tions on Medical Imaging, vol. 36, no. 4, pp. 1027–1036, 2017.

[4] Zhang, M. Ding, F. Meng, and X. Zhang, “Quantitative evaluation of two-factor analysis applied to hepatic perfusion study using contrast1enhanced ultrasound,” IEEE Transactions on Biomedical Engineering, vol. 60, no. 2, pp. 259–267, 2013.

[5] Lueck, T. Kim, P. N. Burns, and A. L. Martel, “Hepatic perfusion imaging using factor analysis of contrast enhanced ultrasound,” IEEE Transactions on Medical Imaging, vol. 27, no. 10, pp. 1449–1457, 2008

[6] L. Guo, D. Wang, Y. Qian, X. Zheng, C. Zhao, X. Li et al., “A two1stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images,” Clinical Hemorheology and Microcirculation, vol. 69, no. 3, pp. 343–354, 2018.

[7] Guo, D. Wang, H. Xu, Y. Qian, C. Wang, X. Zheng et al., “CEUS1based classification of liver tumors with deep canonical correlation anal1ysis and multi-kernel learning,” in Proceedings of the IEEE Engineering in Medicine and Biology Society, EMBC 2017, 2017, pp. 1748–1751.

[8] Wei, C. Zhang, L. Liu, C. Shen, and J. Wu, “Coarse-to-fine: A RNN-based hierarchical attention model for vehicle re-identification,” in Proceedings of the IEEE Conference on Asian Conference on Computer Vision, ACCV 2018, 2018, pp. 575–591.

[9] Rognin, M. Arditi, L. Mercier, P. J. Frinking, M. Schneider, G. Per1renoud et al., “Parametric imaging for characterizing focal liver lesions in contrast-enhanced ultrasound,” IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, no. 11, pp. 2503–2511, 2010.

[10] Turco, P. Frinking, R. Wildeboer, M. Arditi, H. Wijkstra, J. R. Lindner et al., “Contrast-enhanced ultrasound quantification: From kinetic mod1eling to machine learning,” Ultrasound in Medicine Biology, vol. 46, pp. 518–543, 2020

[11] Tran, L. D. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3d convolutional networks,” in Proceedings of the IEEE Conference on International Conference on Computer Vision, ICCV 2015, 2015, pp. 4489–4497.

[12] Wang, C. Lian, D. Yao, D. Zhang, M. Liu, and D. Shen, “Spatial1temporal dependency modeling and network hub detection for functional mri analysis via convolutional-recurrent network,” IEEE Transactions on Biomedical Engineering, vol. 67, no. 8, pp. 2241–2252, 2019.

[13] Zhang, X. He, X. Yu, W. Lu, Z. Zha, and Q. Tian, “A multi-scale spatial-temporal attention model for person re-identification in videos,” IEEE Transactions on Image Processing, vol. 29, pp. 3365–3373, 2020.

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

[14] Wang, R. Girshick, A. Gupta, and K. He, “Non-local neural net1works,” in Proceedings of the IEEE conference on computer vision and pattern recognition, CVPR 2018, 2018, pp. 7794–7803.
[15] Wang and A. Gupta, “Videos as space-time region graphs,” in Proceedings of the Computer Vision European Conference, ECCV 2018, vol. 11209, 2018, pp. 413–431.
[16] Lin, C. Gan, and S. Han, “TSM: temporal shift module for efficient video understanding,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV 2019, 2019, pp. 7082–7092.
[17] Zhang, X. Dai, and Y. Wang, “Dynamic temporal pyramid network: A closer look at multi-scale modeling for activity detection,” in Pro1ceedings of the Asian Conference on Computer Vision, ACCV 2018, vol. 11364, 2018, pp. 712–728.
[18] K. He, X. Zhang, S. Ren, and J. Sun, “Spatial pyramid pooling in deep convolutional networks for visual recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904–1916, 2015
[19] Wang, Y. Xiong, Z. Wang, Y. Qiao, D. Lin, X. Tang et al., “Temporal segment networks for action recognition in videos,” CoRR, vol. abs/1705.02953, 2017.
[20] Zhang, Y.-X. Jiang, J.-B. Liu, M. Yang, Q. Dai, Q.-L. Zhu et al., “Utility of contrast-enhanced ultrasound for evaluation of thyroid nod1ules,” Thyroid, vol. 20, no. 1, pp. 51–57, 2010.
[21] Zhou, P. Zhou, Z. Hu, S. M. Tian, Y. Zhao, W. Liu et al., “Diagnostic efficiency of quantitative contrast-enhanced ultrasound indicators for discriminating benign from malignant solid thyroid nodules,” Journal of Ultrasound in Medicine, vol. 37, no. 2, pp. 425–437, 2018.
[22] He, X. Y. Wang, Q. Hu, X. X. Chen, B. Ling, and H. M. Wei, “Value of contrast-enhanced ultrasound and acoustic radiation force impulse imaging for the differential diagnosis of benign and malignant thyroid nodules,” Frontiers in pharmacology, vol. 9, p. 1363, 2018.
[23] Xuehua, L. Gao, Q. Wu, S. Fang, J. Xu, R. Liu et al., “Differentiation of thyroid nodules difficult to diagnose with contrast-enhanced ultra1sonography and real-time elastography,” Frontiers in oncology, vol. 10, p. 112, 2020.
[24] Luo, Y. Zhang, J. Yuan, X. Yang, L. Pang, L. Ding et al., “Differ1ential diagnosis of thyroid nodules through a combination of multiple ultrasonography techniques: A decision-tree model,” Experimental and therapeutic medicine, vol. 19, no. 6, pp. 3675–3683, 2020.
[25] S. Rinesh, K. Maheswari, B. Arthi, P. Sherubha, A. Vijay et al., “Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms,” Journal of Healthcare Engineering, Vol. 2, pp. 1-9, 2022.
[26] B. Narmadha P. Sherubha, M. Banu chitra, “Multi class feature selection algorithm for breast cancer detection”, International journal of pure and applied mathematics, pp. 301-305, 2018.
[27] P Sherubha, P Amudhavalli, SP Sasirekha, “clone attack detection using random forest and multi-objective cuckoo search classification”, International Conference on Communication and Signal Processing (ICCSP), pp. 0450-0454, 2019.
[28] Peng et al., “A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinson’s disease,” Neurosci. Lett., vol. 651, pp. 88–94, 2017.
[29] Shi et al., “Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset,” Neurocomputing, vol. 194, pp. 87–94, 2016.
[30] Zheng et al., “Improving MRI-based diagnosis of alzheimer’s disease via an ensemble privileged information learning algorithm,” in Proc. 14th Int. Symp. Biomed. Imag., Melbourne, Australia, 2017, pp. 456–459.
[31] S. Hemamalini ,V. D. Ambeth Kumar ,R. Venkatesan,S. Malathi. (2023). Relevance Mapping based CNN model with OSR-FCA Technique for Multi-label DR Classification. Journal of Fusion: Practice and Applications, 11 (2), 90-110.
[32] C. S. Manigandaa,V. D. Ambeth Kumar,G. Ragunath,R. Venkatesan,N. Senthil Kumar. (2023). De-Noising and Segmentation of Medical Images using Neutrophilic Sets. Journal of Fusion: Practice and Applications, 11 (2), 111-123
[33] Sathya Preiya, V., and V. D. Ambeth Kumar. 2023. "Deep Learning-Based Classification and Feature Extraction for Predicting Pathogenesis of Foot Ulcers in Patients with Diabetes" Diagnostics 13, no. 12: 1983. https://doi.org/10.3390/diagnostics13121983

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023

[34] Balakrishnan, Chitra, and V. D. Ambeth Kumar. 2023. "IoT-Enabled Classification of Echocardiogram Images for Cardiovascular Disease Risk Prediction with Pre-Trained Recurrent Convolutional Neural Networks" Diagnostics 13, no. 4: 775. https://doi.org/10.3390/diagnostics13040775.

[35] V. D. Ambeth Kumar,S. Malathi,Abhishek Kumar,Prakash M and Kalyana C. Veluvolu, "Active Volume Control in Smart Phones Based on User Activity and Ambient Noise" ,Sensors 2020, 20(15), 4117; https://doi.org/10.3390/s20154117

[36] N. K. Gupta, A. Sharma, and R. S. Kumar, "A Secure E-Voting System Using Blockchain and Facial Recognition," in *International Journal of Computer Applications*, vol. 182, no. 4, pp. 1-7, 2022.

Doi: https://doi.org/10.54216/FPA.120211
Received: January 28, 2023 Revised: April 21, 2023 Accepted: June 25, 2023