Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FUZZY SYSTEM TO PREDICT PATIENT RISK IN HYPOTHYROID DISEASE USING TSH, T3, T4 HORMONES

MANOHAR ANNAPPA KOLI[1]

[1]Assistant Professor, Department of Computer Science, Karnatak University Dharwad, India

E-mail: [1]koli.manohar@gmail.com

ABSTRACT

The thyroid is an endocrine gland located in the anterior region of the neck: its main task is to produce thyroid hormones (T3 and T4), which controls our entire body. Abnormal production of thyroid hormones can lead to the production of an insufficient or excessive amount of thyroid hormone. Untreated thyroid can lead to serious complications, including heart disease and nerve damage etc. In some cases, it can be fatal. Since patient risk can't to be predicted using discrete values, fuzzy solutions are gaining high importance in healthcare systems. The paper presents novel fuzzy solution to predict patient risk in Hypothyroid using TSH, T3, T4 hormones. The proposed method gives 97.29 percentage of accuracy. Discrete values of risk prediction can't help much in monitoring patient health and disease progress effectively. Using Fuzzy continuous values, it is very easy to predict the progress of patient. Though continuous values based fuzzy systems should not be compared with discrete values and training based machine learning algorithms, still using cut-off values 137 proposed method is compared with existing linear regression, KNN and Bay's algorithms and it is observed proposed method is producing excellent results. The visual results of the proposed method are also justifying the suitability of proposed fuzzy method in hypothyroid risk prediction.

Keywords: *Thyroid Treatment; Hypothyroid; Hyperthyroid; Thyroid Diseases Prediction; Fuzzy Systems; Classifiers; Machine Learning*

1. INTRODUCTION

The thyroid is an endocrine gland found in the neck whose function is to produce hormones Triiodothyronine (T3) and Thyroxine(T4) which it releases into the bloodstream[1]. Thyroid hormone regulates heart rate, body temperature, and above all the body metabolism, that is the way the body uses and consumes nutrients [2]. The thyroid gland may work abnormally and produce high or low thyroid hormones [3]. High thyroid hormone is considered as hyperthyroid disease and low level is considered as hypothyroid disease. Pituitary gland controls the thyroid by releasing thyroid-stimulating hormone (TSH) [4]. According to the American Thyroid Association, there's no cure for hypothyroidism and hyperthyroidism. However, some medications can treat the disease [5, 6]. In literature algorithms are proposed to predict the thyroid related diseases using T3,T4 and TSH hormones using various machine learning algorithms like KNN, Nave Bayes, Support vector machine, and data mining algorithms [7-8]. Some authors tried to predict the disease by using neural network solutions [9, 10].

The proposed methodology is a fuzzy based solution where it produces continuous values, which increases from normal to abnormal range of thyroid. Continuous values help healthcare professionals very much in continuous monitoring of disease.

2. METHODOLOGY

Proposed method includes data fetching, cleaning, analysis, prediction and result analysis as shown in figure 1.

2.1 Data Cleaning and Normalization

In proposed paper Hypothyroid data is collected from the UCI data archive [11] and extracted TSH, T3, T4U (T4) and Outcome values for processing. The normal range of TSH is suggested 0.3–4 mU/l, T3 to be 1.2–2.8 nmol/L and T4 to be 0.8–1.8 ng/dl [12]. The levels of hormones above or below the normal range indicates hyper and hypothyroidism. To improve accuracy of methodology abnormal values of TSH which are above the valid range 7 and below 0 are normalized to 7 and 0 values and then all values are mapped to 0 to 100 scale. All null values are removed before further processing.

2.2 Data Analysis

After normalization to 0 to 100 scales, suggested values of TSH will be 4.28-57.14 mU/l, T3 to be 10.90–26.06 nmol/L and T4 to be 26.57–74.87 ng/dl. After normalization of attributes, frequencies of each feature values (Fij) are calculated and shown in figure 2. And then Confidence of T3, T4 and TSH are calculated using equation 1 and results of confidence of T3, T4 and TSH are shown in figure 3.

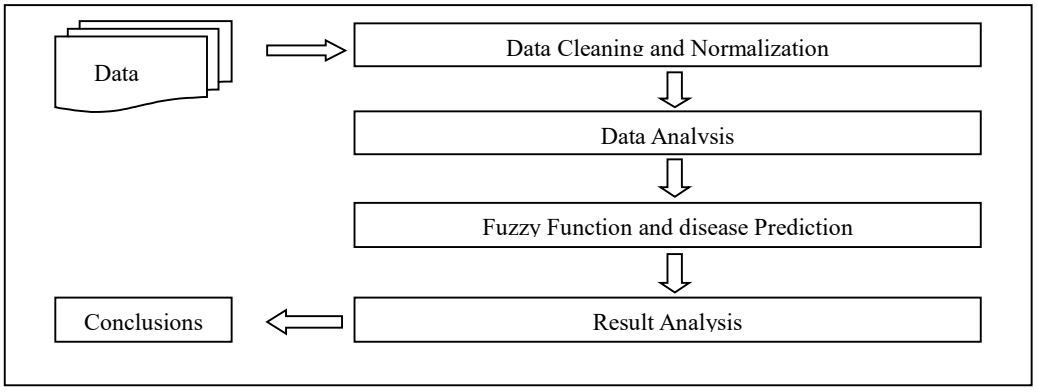

Figure 1 Methodology Of The Paper.

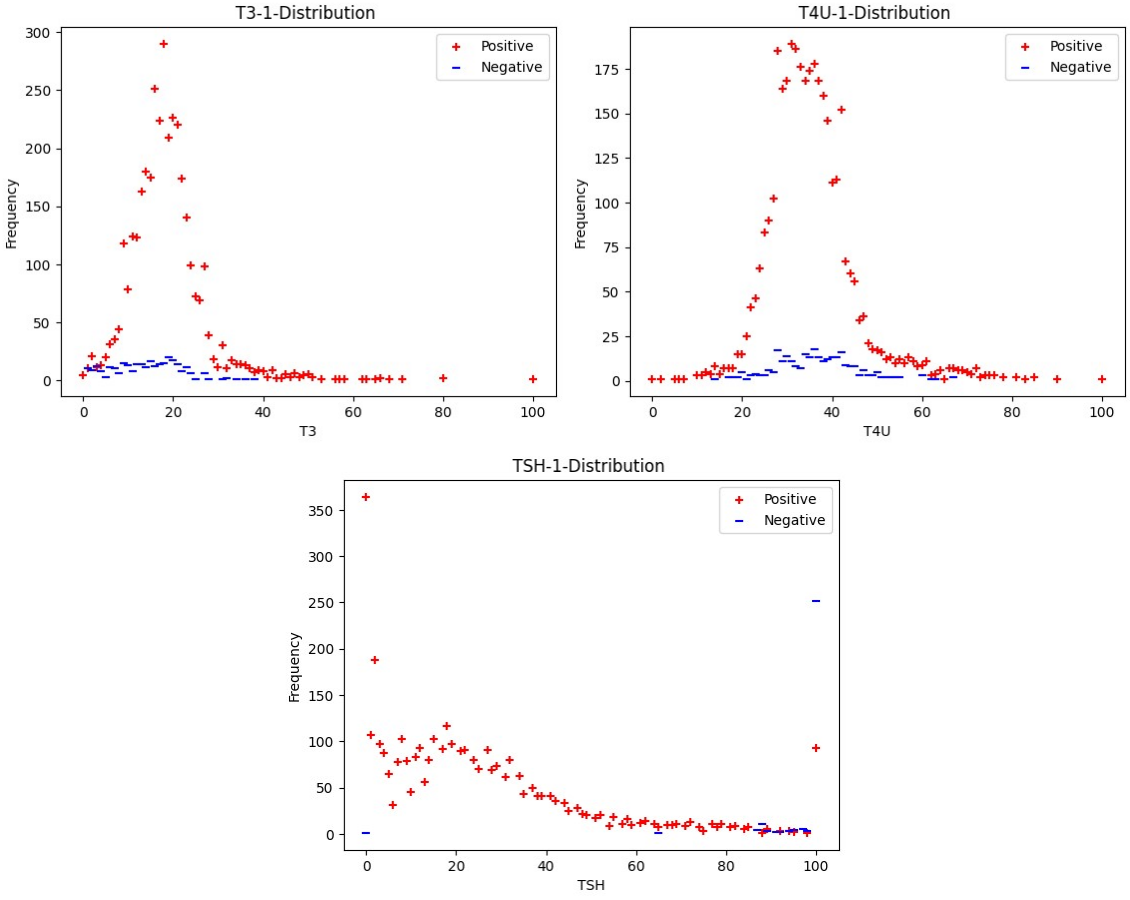

Figure 2 Distribution Of T3, T4 (T4U) And TSH, Frequency Values For + And – Cases.

Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Confidence\ C_{ij} = \left(\frac{(+)\ Frequencies\ of\ Feature\ F_{ij} - (-) Frequencies\ of\ Feature\ F_{ij}}{(+) Frequencies\ of\ Feature\ F_{ij} + (-) Frequencies\ of\ Feature\ F_{ij}}\right) X100 \quad (1)$$

Where i=TSH, T3, T4 and j=0 to100

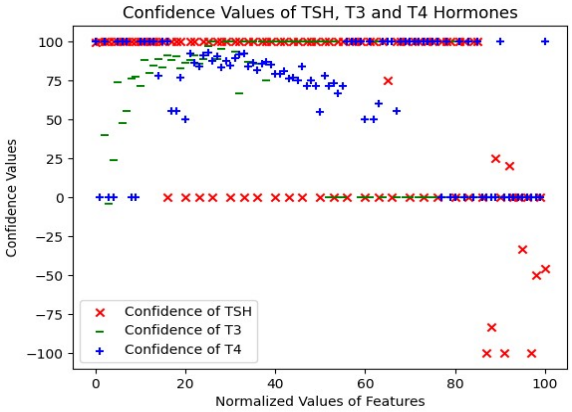

Figure 3 Confidence Values Of TSH, T3 And T4.

2.3 Results Analysis

For all dataset rows, Fuzzy function F(x) is calculated using equation 2 and plotted in figure 4. Results of F(x) shows all blue colored negative values (Normal peoples) are moved towards left side of the figure and all red colored positive values (Hypothyroid Peoples) are moved towards the right side of the figure 4. The distribution of results F(x) in figure 4 clearly shows left to right increase of F(x) also increases the frequency of red colored positive cases.

To finalize the fuzzy results Z1 value is calculated using fuzzy equation 3. Where Z1 value indicates severity of disease, value 1indicates 100% Hypothyroid and 0 indicates 100% normal thyroid cases. In between 0 to 1 higher values indicates more severity and lower values indicates low severity of disease. To calculate fuzzy Z value constants k1=58 and k2=261 are used.

Discrete values of risk prediction can't help in monitoring patient health and disease progress. Using Fuzzy continuous values it is very easy to compare the progress of patient. Though continuous values based fuzzy systems should not be compared with discrete values and training based machine learning algorithms, still using cut-off 137, proposed method is compared with existing linear regression (LR), KNN and Bay's algorithms and it is observed that the proposed method is giving excellent results as shown in table 1 and also visualized in figure 5.

$$F(x) = \sum_{i=0}^{i=len(Features)-1} Confidence\ of\ C_{ij} \quad (2)$$

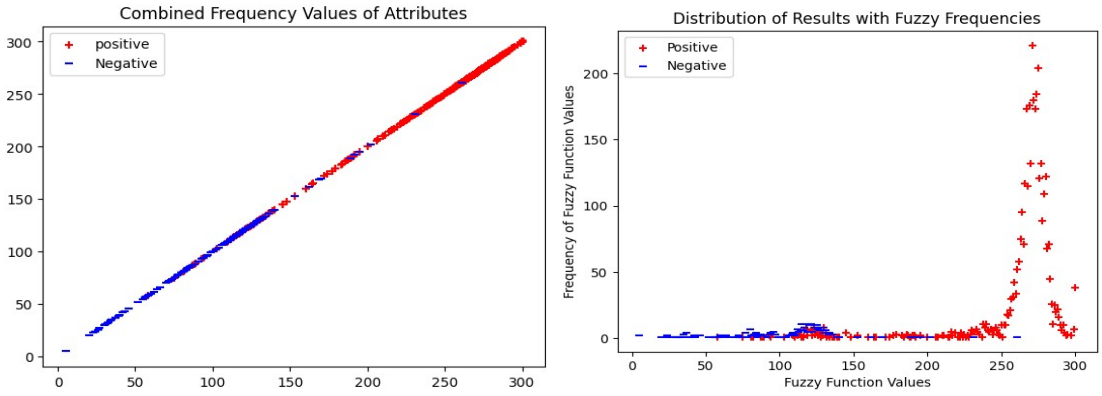

Figure 4 Distributions Of F(X) Frequencies For + And - Cases.

Journal of Theoretical and Applied Information Technology
31st March 2023. Vol.101. No 6

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Z1 = \begin{cases} 0 & Frequency\ of\ F(x) < k1 \\ \frac{(Frequency\ of\ F(x) - k1)}{(k2 - k1)} & k1 \le Frequency\ of\ F(x) \le k2 \\ 1 & Frequency\ of\ F(x) > k2 \end{cases} \quad Where\ k1{=}58 \text{ and } K2{=}261 \qquad (3)$$

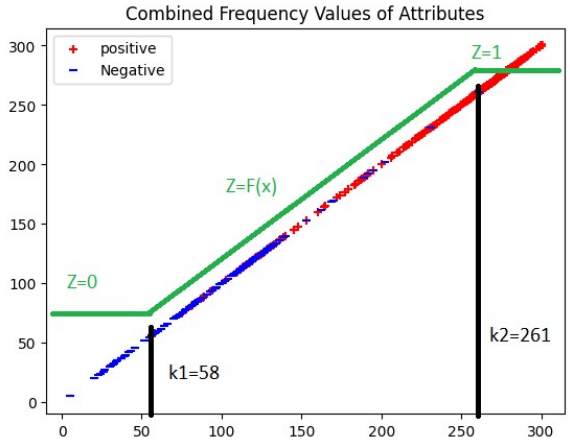

Figure 5 Distributions Of F(X) Frequencies For + And – Cases With Fuzzy Cutoff Values.

$$Z2 = \begin{cases} Positive(1),\ F(x) > 137 \\ Negative(0),\ F(x) \le 137 \end{cases} \qquad (4)$$

Table 1 Comparison Of Proposed Methodology With Existing Popular Classification Methods.

S.N	Methodology	False Negative (FN)	False Positive (FP)	True Negative (TN)	True Positive (TP)	Accuracy (ACC)
1	Proposed Method(PM) (Cutoff=-136)	279	12	90	3391	97.29
2	Linear regression(LR)	237	79	54	3402	92.28
3	KNN(KNN)	279	105	12	3376	96.89
4	Naive Beyes(NB)	289	118	2	3363	96.81

3. CONCLUSIONS

The thyroid gland is considered as a power house of body, detecting normal and abnormal working of thyroid such as hypo and hyper thyroidism is very import. This paper has presented novel fuzzy solution for predicting patient risk in hypothyroid using TSH, T3, T4 hormones with an accuracy of 97.29 percentages. In comparative study, the proposed method is showing better results than LR, KNN, and NB. The visual results of the proposed method are also justifying the suitability of proposed method in hypothyroid risk prediction.

REFERENCES:

[1] Lerina Aversanoa, Mario Luca Bernardia , Marta Cimitileb, Martina Iammarinoa , Paolo Emidio Macchiac , Immacolata Cristina Nettorec , Chiara Verdonea.Thyroid Disease Treatment prediction with machine learning approaches. 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Procedia Computer Science 192 (2021) 1031–1040.

[2] Saffron Whitehead. Endocrinology: An integrated approach. 01 2001.

[3] Hamilton Romaldini, Jos Augusto Sgarbi, and Chady Satt Farah. Subclinical thyroid disease: subclinical hypothyroidism and hyperthyroidism. Arquivos brasileiros de endocrinologia e metabologia, 48(1):147158, February 2004

[4] Zadeh, L. A. (1965). Fuzzy sets, information, and control. vol, 8, 338-353.

[5] Watson, T. (2008). CauseWired: plugging in, getting involved, and changing the world. John Wiley & Sons.

[6] GRANGER, J. Annual crime report shows a rise in assaults.

[7] A. Begum and A. Parkavi. Prediction of thyroid disease using data mining techniques. In 2019 5th International Conference on Advanced Computing Communication Systems (ICACCS), pages 342–345, 2019.

[8] Irina Ionit and Liviu Ionit. Prediction of thyroid disease using data mining techniques. BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 7(3), 2016.

[9] Shiva Borzouei, Hossein Mahjub, NegarAsaad Sajadi, and Maryam Farhadian. Diagnosing thyroid disorders: Comparison of logistic regression and neural network models. Journal of Family Medicine and Primary Care, 9:1470, 06 2020.

[10] Mahdiyar O. Obeidavi M. R, Rafiee A. Diagnosing thyroid disease by neural networks. 2017.

[11] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[12] Kinjo Y, Takasu N, Komiya I, Tomoyose T, Takara M, Kouki T, Shimajiri Y, Yabiku K, Yoshimura H. Remission of Graves' hyperthyroidism and A/G polymorphism at position 49 in exon 1 of cytotoxic T-lymphocyte-associated molecule-4 gene. J Clin Endocrinol Metab. 2002; 87(6):2593–2596. doi: 10.1210/jc.87.6.2593.