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

MANOHAR ANNAPPA KOLI

1Assistant Professor, Department of Computer Science, Karnatak University Dharwad, India
E-mail: 1koli.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.
Where \( i = \text{TSH}, \text{T3}, \text{T4} \) and \( j = 0 \) to 100

\[
\text{Confidence } C_{ij} = \left( \frac{(+ \text{ Frequencies of Feature } F_i) - (- \text{ Frequencies of Feature } F_i)}{(+ \text{ Frequencies of Feature } F_i) + (- \text{ Frequencies of Feature } F_i)} \right) \times 100
\]

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 \( Z_1 \) value is calculated using fuzzy equation 3. Where \( Z_1 \) value indicates severity of disease, value 1 indicates 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 \( k_1 = 58 \) and \( k_2 = 261 \) 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.

\[
F(x) = \sum_{i=0}^{i=\text{len(Features)}-1} \text{Confidence of } C_{ij}
\]
Where $k_1 = 58$ and $k_2 = 261$

\[ Z_1 = \begin{cases} 0 & \text{Frequency of } F(x) < k_1 \\ \frac{(Frequency \text{ of } F(x) - k_1)}{(k_2 - k_1)} & \text{Frequency of } F(x) \leq k_2 \\ 1 & \text{Frequency of } F(x) > k_2 \end{cases} \] (3)

\[ Z_2 = \begin{cases} \text{Positive(1), } F(x) > 137 \\ \text{Negative(0), } F(x) \leq 137 \end{cases} \] (4)

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

<table>
<thead>
<tr>
<th>S.N</th>
<th>Methodology</th>
<th>False Negative (FN)</th>
<th>False Positive (FP)</th>
<th>True Negative (TN)</th>
<th>True Positive (TP)</th>
<th>Positive (ACC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proposed Method (PM) (Cutoff = 136)</td>
<td>279</td>
<td>12</td>
<td>90</td>
<td>3391</td>
<td>97.29</td>
</tr>
<tr>
<td>2</td>
<td>Linear regression (LR)</td>
<td>237</td>
<td>79</td>
<td>54</td>
<td>3402</td>
<td>92.28</td>
</tr>
<tr>
<td>3</td>
<td>KNN(KNN)</td>
<td>279</td>
<td>105</td>
<td>12</td>
<td>3376</td>
<td>96.89</td>
</tr>
<tr>
<td>4</td>
<td>Naive Bayes (NB)</td>
<td>289</td>
<td>118</td>
<td>2</td>
<td>3363</td>
<td>96.81</td>
</tr>
</tbody>
</table>

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 thyroïdism is very important. 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:


