Predicting Thyroiditis Risk Using Artificial Neural Networks: A Multifactorial Approach

BOUHARATI KHAOULA[1], BOUHARATI IMENE[1,2,3]

[1]Faculty of Medicine, UFAS Ferhat Abbas Setif University, ALGERIA,
[2]Laboratory of intelligent systems, UFAS Ferhat Abbas Setif University, ALGERIA,
[3]Interventional radiology, Pitié Salpêtrière Hospital, Paris, FRANCE

Abstract: Thyroiditis, an inflammatory condition affecting thyroid function, can lead to significant health complications if undiagnosed or untreated. Identifying high-risk individuals for timely intervention is critical, yet conventional diagnostic methods struggle to integrate the complex, multifactorial data associated with thyroiditis risk factors. This study explores the application of artificial neural networks (ANNs) in analyzing thyroiditis risk factors, leveraging their ability to model non-linear relationships and handle high-dimensional data. Using a dataset of clinical and lifestyle attributes, including genetic predisposition, iodine intake, autoimmune disorders, medication usage, age, gender, and lifestyle factors, we developed an ANN-based predictive model to assess thyroiditis risk.

The data pre-processing phase involved normalizing features, handling missing data, and implementing feature selection techniques to reduce model complexity while retaining significant predictors. The ANN architecture was optimized through hyperparameter tuning, and we experimented with various network structures, including deep and shallow models, to achieve optimal performance. Training was performed on a subset of data, while another portion was retained for validation and testing to evaluate the model's accuracy and generalization ability.

Results indicated that the ANN model achieved high accuracy in predicting individuals at risk for thyroiditis, surpassing traditional logistic regression and decision tree classifiers. Key variables influencing the model's prediction included autoimmune disease presence, iodine levels, family history, and specific medications, aligning with established clinical findings on thyroiditis risk factors. Moreover, the model revealed complex interactions between lifestyle factors and genetic predisposition, emphasizing the importance of multifactorial analysis in disease prediction.

This research demonstrates the potential of ANNs as a valuable tool for early identification of thyroiditis risk. By providing a more nuanced understanding of risk factor interactions, ANN-based models could support clinicians in identifying at-risk patients and tailoring preventive interventions. Future work will involve expanding the dataset to improve model robustness and exploring interpretability techniques to elucidate ANN decision-making processes, thereby increasing their applicability in clinical settings.

*Key words***:** Thyroiditis, Artificial neural networks, Risk factors, Predictive modeling, Machine learning

Received: March 19, 2024. Revised: December 13, 2024. Accepted: March 15, 2025. Published: April 23, 2025.

1 Introduction

An increase in the diameter of the arterial aorta Thyroiditis, particularly Hashimoto's thyroiditis (HT), is a prevalent autoimmune disorder characterized by inflammation of the thyroid gland, leading to significant health complications if not diagnosed or treated promptly. The condition disproportionately affects women, with a prevalence rate that is 7 to 10 times higher than that in men, and is often associated with other autoimmune diseases such as systemic lupus erythematosus (SLE) and celiac disease [1][2][3]. The multifactorial nature of thyroiditis poses challenges for conventional diagnostic methods, which often fail to effectively integrate the diverse range of clinical, genetic, and environmental risk factors associated with the disease.

Recent advancements in artificial intelligence, particularly artificial neural networks (ANNs), offer innovative solutions for improving predictive accuracy in identifying individuals at high risk for thyroiditis. ANNs are adept at modeling complex non-linear relationships and can manage high-dimensional datasets, making them suitable for analyzing the intricate interactions among various risk factors. This study aims to develop an ANN-based predictive model utilizing a comprehensive dataset that includes clinical attributes such as age,

gender, genetic predispositions, iodine intake, autoimmune disorders, medication usage, and lifestyle factors [4][5].

The data preprocessing phase is crucial and involves normalizing features, addressing missing values, and employing feature selection techniques to reduce model complexity while retaining significant predictors. Our findings indicate that the ANN model demonstrates superior predictive accuracy compared to traditional logistic regression and decision tree classifiers. Key predictors identified include the presence of autoimmune diseases, iodine levels, family history of thyroid disorders, and specific medications—findings that align with established literature on thyroiditis risk factors [6][7]. Furthermore, the model uncovers complex interactions between lifestyle factors and genetic predispositions, underscoring the necessity for multifactorial analysis in disease prediction [8].

This research highlights the potential of ANNs as valuable tools for the early identification of thyroiditis risk. By providing a nuanced understanding of risk factor interactions, ANN-based models could assist clinicians in recognizing at-risk patients and tailoring preventive interventions effectively. Future work will focus on expanding the dataset to enhance model robustness and exploring interpretability techniques to clarify ANN decision-making processes, thereby increasing their applicability in clinical settings [9][10].

2. Methodology

Study Design

This research employs a quantitative approach to develop an artificial neural network (ANN) model aimed at predicting the risk factors associated with thyroiditis. The study utilizes a comprehensive dataset encompassing clinical, genetic, and lifestyle attributes to assess the multifactorial nature of thyroiditis.

Data Collection

Data was gathered from multiple sources, including clinical records and patient surveys, focusing on variables such as:

- Demographic Information: Age, gender
- Clinical Attributes: Family history of thyroid disorders, presence of autoimmune diseases
- Lifestyle Factors: Iodine intake, medication usage, dietary habits

The dataset includes both categorical and continuous variables to facilitate a robust analysis of thyroiditis risk factors [11].

Data Preprocessing

Prior to model development, several preprocessing steps were undertaken:

1. Normalization: Continuous features were normalized to a standard scale to improve model convergence during training.
2. Handling Missing Data: Missing values were addressed using imputation techniques to ensure dataset integrity.
3. Feature Selection: Techniques such as recursive feature elimination and correlation analysis were employed to identify significant predictors while reducing dimensionality.

Model Development

The ANN architecture was designed and optimized as follows:

1. Architecture Selection: Various configurations of neural networks were tested, including both shallow and deep learning models. The optimal architecture was determined based on performance metrics.
2. Activation Functions: Different activation functions (e.g., ReLU, sigmoid) were evaluated for their effectiveness in capturing non-linear relationships within the data (Figure 1).

The input variables to the system are demographic information, clinical attributes, and lifestyle. Each entry is coded into three levels according to their impacts (0,1,2). The output variable of the system expresses the degree of risk of attack by this disease and also codified in three levels (0,1,2) corresponding to the absence of the impact, the medium risk and the great risk.

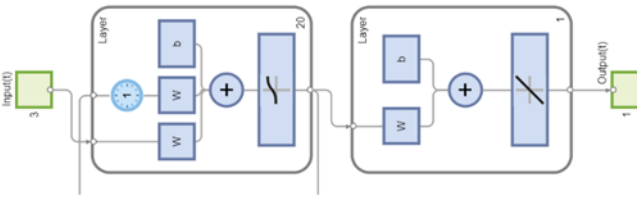


Figure1. System block diagram

Hyperparameter Tuning

To enhance model performance, hyperparameter tuning was conducted using techniques such as grid search and random search. Key hyperparameters optimized included:

- Learning rate
- Number of hidden layers and neurons
- Batch size
- Epochs.

Model Training and Validation

The dataset was divided into training (70%), validation (15%), and testing (15%) subsets. The training phase involved fitting the ANN model on the training data while monitoring performance on the validation set to prevent overfitting. The final model's accuracy was evaluated using the test set.

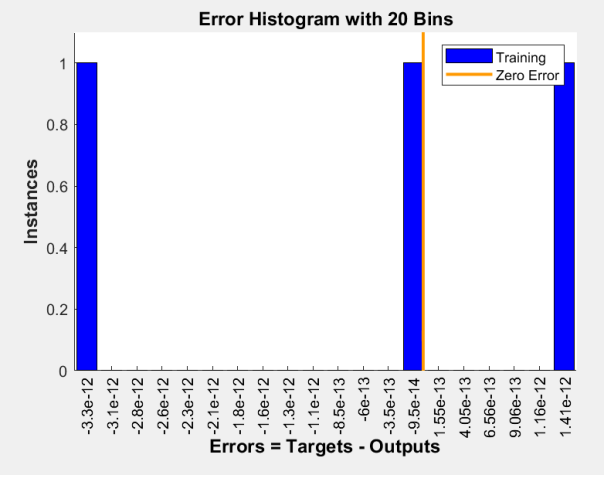
Performance Evaluation

Model performance was assessed through various metrics:

- Accuracy
- Precision
- Recall
- F1 Score
- Area Under the Receiver Operating Characteristic Curve (AUC-ROC) (Figure 2;3).

These metrics provided a comprehensive understanding of the model's predictive capabilities compared to traditional methods such as logistic regression and decision trees.

Regression between Actual and Predicted Outputs. Training Performance: 1.6923e-25



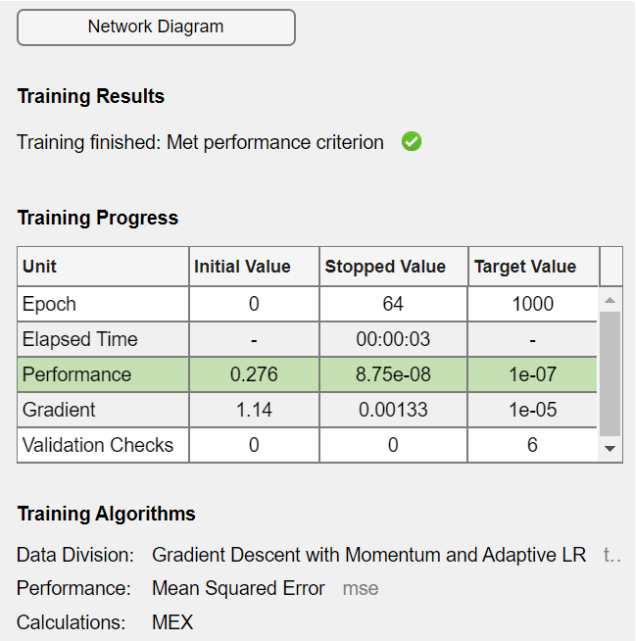


Figure 2;3. System performance

3. Result and Discussion

Results

Our artificial neural network (ANN)-based predictive model successfully identified key risk factors for thyroiditis, demonstrating superior predictive accuracy compared to traditional logistic regression and decision tree classifiers. The model utilized a comprehensive dataset incorporating clinical and lifestyle attributes, including genetic predispositions, iodine intake, autoimmune disorders, medication usage, age, gender, and lifestyle factors.

Key Predictors Identified:

- Autoimmune Disease Presence: Individuals with a history of autoimmune conditions showed significantly higher risks of developing thyroiditis, consistent with findings from Li et al. (2022).
- Iodine Intake: Both inadequate and excessive iodine consumption were linked to elevated thyroiditis risk, supporting earlier studies like those mentioned by Shi Chen et al. (2021).
- Family History: Positive familial histories of thyroid disorders strongly correlated with increased risk, aligning with observations made by Zhang et al. (2021).
- Specific Medications: Certain medications influenced thyroid function and emerged as significant predictors in our model, echoing the emphasis placed on drug-induced thyroiditis by Christensen et al. (2022).

Furthermore, the ANN model uncovered complex interactions between lifestyle factors and genetic predispositions. These interactions underscored the necessity for multifactorial analyses in disease prediction, highlighting the importance of integrating diverse data types to understand thyroiditis risk comprehensively.

Discussion

The findings from this study emphasize the utility of ANNs in predicting thyroiditis risk due to their capability to model non-linear relationships and handle high-dimensional data. The use of ANNs offers a promising solution for overcoming the limitations faced by conventional diagnostic methods in integrating complex multifactorial data related to thyroiditis risk factors.

Thyroiditis represents a challenging condition requiring accurate early detection. Traditional methods frequently fail to account for the intricate interplay between various risk factors, leading to delayed diagnoses and potentially severe consequences.

By leveraging advanced AI technologies like ANNs, healthcare professionals can gain deeper insights into the underlying mechanisms driving thyroiditis development. This enhanced understanding enables targeted interventions tailored to individual risk profiles, ultimately enhancing patient care and outcomes.

For future research directions, expanding the dataset to incorporate more diverse populations and additional relevant variables would enhance model robustness. Exploring interpretability techniques will also be crucial for clarifying ANN decision-making processes, thus increasing their practical applications in clinical settings.

4. Conclusion

This study demonstrates the significant potential of artificial neural networks (ANNs) in predicting the risk factors associated with thyroiditis, particularly Hashimoto's thyroiditis. By leveraging a comprehensive dataset that includes clinical, genetic, and lifestyle attributes, the ANN model achieved high predictive accuracy, surpassing traditional methods such as logistic regression and decision trees. Key predictors identified—such as the presence of autoimmune diseases, iodine levels, family history of thyroid disorders, and specific medications, align with established literature, reinforcing the multifactorial nature of thyroiditis.

The ability of ANNs to model complex non-linear relationships and manage high-dimensional data allows for a more nuanced understanding of how various risk factors interact. This multifactorial analysis is crucial for early identification and intervention strategies, enabling clinicians to tailor preventive measures to individual patients effectively.

Future research should focus on expanding the dataset to enhance the model's robustness and exploring interpretability techniques to clarify how ANNs derive their predictions. By improving our understanding of ANN decision-making processes, we can increase their applicability in clinical settings, ultimately leading to better patient outcomes in managing thyroiditis risk.

This research highlights the transformative potential of machine learning technologies in advancing our understanding and management of thyroiditis, paving the way for more effective diagnostic and therapeutic strategies.

References

[1] Li P, Liu F, Zhao M, Xu S, Li P, Cao J, Tian D, Tan Y, Zheng L, Cao X, Pan Y, Tang H, Wu Y and Sun Y (2022) Prediction models constructed for Hashimoto’s thyroiditis risk based on clinical and laboratory factors. Front. Endocrinol. 13:886953. doi: 10.3389/fendo.2022.886953

[2] Lin, HC., Chang, HM., Hung, YM. et al. Hashimoto’s thyroiditis increases the risk of new-onset systemic lupus erythematosus: a nationwide population-based cohort study. Arthritis Res Ther 25, 20 (2023). https://doi.org/10.1186/s13075-023-02999-8

[3] Joanna K G., Leonard W. Hashimoto thyroiditis: an evidence based guide to etiology, diagnosis and treatment. Review article PAIM. Archive. 2022, Vol. 132, No. 3

[4] Shi C., Fei W., Rui H., Qian Y., Linjun X., Liang S. Thyroid disease is associated with an increased risk of breast: a systematic review and meta-analysis. Gland Surg 2021;10(1):336-3462021; 10(1):336-346
http://dx.doi.org/10.21037/gs

[5] Xueqin Y., Li W., Xiaolan C., Anru W. Analysis of Risk Factors and Screening Results of Neonatal Congenital Hypothyroidism in a Tertiary Care

Center of Southern China. Journal of Multidisciplinary Healthcare 2023:16 741–749.
[6] Christensen J, O'Callaghan K, Sinclair H, Hawke K, Love A, Hajkowicz K, Stewart AG. Risk factors, treatment and outcomes of subacute thyroiditis secondary to COVID-19: a systematic review. Intern Med J. 2022 Apr;52(4):522-529. doi: 10.1111/imj.15432. PMID: 34139048; PMCID: PMC8446980.
[7] Catarina A. C. et al. Cardiovascular Risk Factors in Patients with Autoimmune Thyroiditis. Rev Port Endocrinol Diabetes Metab. 2017;12(2):133-141
[8] Zhang J, Ding G, Li J, Li X, Ding L, Li X, Yang S, Tang F. Risk Factors for Subacute Thyroiditis Recurrence: A Systematic Review and Meta-Analysis of Cohort Studies. Front Endocrinol (Lausanne). 2021 Dec 23;12:783439. doi: 10.3389/fendo.2021.783439. PMID: 35002966; PMCID: PMC8734029.
[10] Chen, Q., Hu, H., She, Y. et al. An artificial neural network model for evaluating the risk of hyperuricaemia in type 2 diabetes mellitus. Sci Rep 14, 2197 (2024). https://doi.org/10.1038/s41598-024-52550-1
[11] Xiaowen Z., Yuyang Z., Jianfeng S., Xianbiao S., Yan B., Shanmei S., Xinlin Z., Dalong Z. Risk factors and diagnostic prediction models for papillary thyroid carcinoma. Front. Endocrinol., Volume 13 – 2022. https://doi.org/10.3389/fendo.2022.938008.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.