Journal of
Modern Medical Information Sciences

Autumn 2023. Vol 9. Issue 3

Research Paper

Presenting a Smart Model for Distinguishing Patients With Thyroid Diseases From Healthy People by Combining Particle Swarm Optimization Algorithm and Artificial Neural Network

***Mohammadjavad Hosseinpoor**[1]

1. Department of Computer Engineering, Faculty of Computer Engineering, Estahban Branch, Islamic Azad University, Estahban, Iran.

Use your device to scan and read the article online

Citation Hosseinpoor MJ. [IPresenting a smart Model for Distinguishing Patients With Thyroid Diseases From Healthy People by Combining Particle Swarm Optimization Algorithm and Artificial Neural Network (Persian)]. Journal of Modern Medical Information. 2023; 9(3):222-233. https://doi.org/10.32598/JMIS.9.3.6

https://doi.org/10.32598/JMIS.9.3.6

CC BY NC

Article Info:

Received: 04 Apr 2023

Accepted: 07 Aug 2023

Available Online: 01 Oct 2023

Key words:
Prediction, Thyroid disease, Disease diagnosis, Evolutionary algorithm, Learning model

ABSTRACT

Objective Thyroid diseases are common disorders worldwide. The timely diagnosis and control of this disease can prevent its progression and reduce associated complications. This study proposes a novel hybrid method by combining particle swarm optimization (PSO) algorithm and artificial neural network (ANN) for the timely detection of thyroid disorders.

Methods This is an applied survey study, conducted in 2022. In this study, the target population consisted of the data of 400 patients referred to Imam Reza Hospital in Lar County, Iran from 2021 to 2022 which were collected by field study. Among them, 300 had thyroid disease and 100 were healthy. MATLAB software was used for implementing the proposed learning model and analyzing the results.

Results The regression coefficient of the proposed model in there modes of training, validation, and testing were 0.98, 0.97, and 0.95, respectively. The area under the ROC curve was 0.98, the error rate was 0.004, and the overall accuracy was 96%.

Conclusion The proposed model can distinguish patients with thyroid disease from healthy individuals with acceptable accuracy and low errors. This model can be used as a useful tool in predicting thyroid diseases.

* *Corresponding Author:*
Mohammadjavad Hosseinpoor, Assistance professor.
Address: Department of Computer Engineering, Faculty of Computer Engineering, Estahban Branch, Islamic Azad University, Estahban, Iran.
Tel: +98 (910) 3076959
E-mail: mj.hosseinpoor@iau.ac.ir

Journal of
Modern Medical Information Sciences

Extended Abstract

Introduction

Thyroid diseases are common disorders worldwide. Various studies have shown their high prevalence in different countries. The timely diagnosis of these diseases and their control can prevent their progress and reduce their consequences. The use of appropriate methods based on artificial intelligence and learning machine algorithms can be useful to make a correct diagnosis in medical cases. The use of smart techniques can provide results that show the factors of various diseases, and based on these results, physicians can act to prevent the disease. Extracting knowledge from massive amounts of data using learning models can lead to the identification of laws governing diseases and provide valuable information to health professionals to identify the causes of diseases, diagnose, predict, and treat patients according to environmental factors. In this research, we aim to use learning models based on artificial neural network (ANN) to predict thyroid diseases.

Methods

This is an applied survey study that was conducted in 2022. The data of more than 400 patients referred to Imam Reza Hospital in Lar County, Iran from 2021 to 2022 were examined. The proposed method is a hybrid algorithm resulting from the combination of particle swarm optimization (PSO) algorithm and ANN. The process includes three main phases. First, the dataset of thyroid patients is used for pre-processing. In this phase, the input dataset are normalized, and incomplete records are removed from the data set. Then, the normalized dataset enters the second phase for clustering and data labeling. In this phase, the K-means algorithm is used to create clusters in the PSO algorithm. Then, the clusters enter the PSO algorithm to find the optimal clusters. After performing the optimization operation by the PSO algorithm, the optimal clusters containing two clusters of healthy people and patients (labeled datasets) are entered into the next phase as output. In the third phase, the labeled input dataset enters the neural network model designed for the learning process of the intelligent model. The neural network used in this phase was the feed-forward neural network with one input layer, 20 hidden layers, and one output layer. After the training of the intelligent model, this model was used to distinguish people with thyroid diseases from healthy people.

Results

The proposed model was trained on a dataset with 400 samples, including 300 people with thyroid diseases and 100 healthy people. The regression coefficient (R) value of the proposed model in three modes of training, validation, and test was 0.98, 0.97 and 0.95, respectively. The area under the ROC curve was 0.98, and the error rate was 0.004. The results of the proposed method were compared with previous methods based on accuracy. According to Table 1, the proposed model was able to predict thyroid disease with 96% accuracy, compared to previous methods.

Conclusion

In this study, an intelligent method based on PSO algorithm and ANN was presented for distinguishing patients with thyroid diseases from healthy people. The proposed model's goal was to improve the process of thyroid disease detection. The main part this model was ANN which tried to find the hidden patterns of patients compared to healthy people by using the labeled datasets. Since the way the ANN works depends on how its structure is defined, we used the PSO algorithm is proposed.

Table 1. Comparing the proposed method with other methods in terms of accuracy

Reference	Method	Accuracy (%)
[16]	Neural fuzzy classification	92
[18]	backpropagation algorithm + ANN	94
[19]	ANN	82
[20]	CNN	94.4
Current study	PSO algorithm + ANN	96

Ethical Considerations

Compliance with ethical guidelines

This study has ethical approval from the Islamic Azad University of Estahban Branch (Code: IR.IAUESTAHBAN.REC.1401.023).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

The authors declared no conflict of interest.

مقاله پژوهشی

ارائه مدلی هوشمند برای تشخیص بیماران مبتلا به بیماری های تیروئید از افراد سالم با ترکیب الگوریتم بهینه سازی ازدحام ذرات و شبکه عصبی مصنوعی

*محمدجواد حسین‌پور

۱. گروه مهندسی کامپیوتر، دانشکده مهندسی کامپیوتر، واحد استهبان، دانشگاه آزاد اسلامی، استهبان، ایران.

Use your device to scan and read the article online

Citation Hosseinpoor MJ. [Intelligent Prediction of Thyroid Disease Using an Algorithm Based on an Artificial Neural Network (Persian)]. Journal of Modern Medical Information. 2023; 9(3):222-233. https://doi.org/10.32598/JMIS.9.3.6

https://doi.org/10.32598/JMIS.9.3.6

اطلاعات مقاله:

تاریخ دریافت: ۱۵ فروردین ۱۴۰۲

تاریخ پذیرش: ۱۶ مرداد ۱۴۰۲

تاریخ انتشار: ۰۹ مهر ۱۴۰۲

چکیده

هدف بیماری‌های تیروئید در سراسر جهان گسترده شده است. مطالعات تحقیقاتی مختلف نشان می‌دهد تعداد زیادی از افراد در جوامع مختلف به این بیماری دچار می‌شوند. همچنین، تشخیص به موقع این بیماری و کنترل آن می‌تواند جلوی پیشرفت آن را بگیرد و پیامدهای ناشی از آن را کاهش دهد. در این راستا، مطالعه پیش‌رو یک الگوریتم ترکیبی تکاملی حاصل از آمیختگی الگوریتم بهینه‌سازی ازدحام ذرات و شبکه عصبی مصنوعی جهت تشخیص به موقع این بیماری ارائه کرده است.

روش ها پژوهش حاضر از نوع کاربردی‌پیمایشی است که در سال ۱۴۰۱ انجام شده است. در اینجا از روش مجموعه داده‌های اولیه برای جمع‌آوری داده‌ها استفاده شد. جامعه آماری موردنظر شامل ۴۰۰ مورد اطلاعات ثبت‌شده بیماران از سال ۱۴۰۰ تا ۱۴۰۱ در تحقیقی میدانی از افراد مراجعه‌کننده به بیمارستان امام رضا (ع) شهرستان لارستان است. از این میان، ۳۰۰ نفر دارای بیماری تیروئید و ۱۰۰ نفر سالم بودند. در این پژوهش برای پیاده‌سازی مدل یادگیری پیشنهادی و همچنین تجزیه‌وتحلیل و بررسی نتایج از نرم‌افزار متلب استفاده شده است.

یافته ها نتایج نشان داد، ضریب رگرسیون مدل پیشنهادی در ۳ حالت آموزش، اعتبارسنجی و تست به‌ترتیب دارای مقادیر (۰/۹۸، ۰/۹۷ و ۰/۹۵)، منحنی راک برابر با ۰/۹۸، میزان خطا برابر با ۰/۰۰۴ و دقت کل سیستم برابر با ۹۶ درصد می‌باشد.

نتیجه گیری باتوجه‌به نتایج حاصله، مدل پیشنهادی می‌تواند با دقت قابل قبولی، پیش‌بینی بیماری تیروئید در افراد را انجام دهد و باعث کاهش میزان اشتباه شود. همچنین از این مدل می‌توان به‌عنوان یک ابزار مفید در پیش‌بینی تیروئید به کار برده شود.

کلیدواژه‌ها:

پیش‌بینی، بیماری تیروئید، تشخیص بیماری، الگوریتم تکاملی، مدل یادگیر

* نویسنده مسئول:

دکتر محمدجواد حسین‌پور

نشانی: استهبان، دانشگاه آزاد اسلامی، واحد استهبان، دانشکده مهندسی کامپیوتر، گروه مهندسی کامپیوتر.

تلفن: ۳۰۷۶۹۵۹ (۹۱۰) ۹۸+

پست الکترونیکی: mj.hosseinpoor@iau.ac.ir

مقدمه

پیش‌بینی بیماری‌ها یک فرآیند بسیار سخت و پیچیده برای افراد خبره در علوم پزشکی محسوب می‌شود [۱]، بنابراین، لزوم استفاده از روش‌های مناسب هوش مصنوعی و مدل‌های یادگیر جهت اتخاذ تشخیص صحیح در مسائل پزشکی ضروری به نظر می‌رسد [۲]. درنتیجه استفاده از الگوریتم‌ها و تکنیک‌های هوشمند و یادگیر می‌تواند الگوها و نتایجی را برای محققین فراهم کند [۳] که عوامل ابتلا به بیماری‌های مختلف را نشان دهد و براساس آن نتایج، پزشکان و دست‌اندرکاران علوم پزشکی در پیشگیری آن بیماری اقدام کنند [۴].

استخراج دانش از میان حجم انبوه داده‌ها بااستفاده‌از مدل‌های یادگیر می‌تواند منجر به شناسایی قوانین حاکم بر بیماری‌ها شود و اطلاعات ارزشمندی را به منظور شناسایی علل بیماری‌ها، تشخیص، پیش‌بینی و درمان بیماری‌ها باتوجه‌به عوامل محیطی در اختیار متخصصان حوزه سلامت قرار دهد [۵] و این مسئله به معنای عمر و ایجاد آرامش برای افراد جامعه است [۶]. در این پژوهش، از مدل‌های یادگیر جهت پیش‌بینی بیماری تیروئید استفاده شده است. هورمون‌های تیروئید پس از ساخته شدن در غده تیروئید به جریان خون منتقل می‌شوند و بسیاری از فرآیندهای شیمیایی را در بخش‌های مختلف بدن کنترل می‌کنند [۷]. این هورمون‌ها برای تکامل و عملکرد طبیعی مغز و سیستم عصبی و برای حفظ دمای بدن و انرژی ضروری هستند [۸]. غده تیروئید هورمون‌های خود را در خون ترشح می‌کند [۹].

این هورمون‌ها(یعنی تری‌یدوتیرونین و تیروکسین)واکنش‌های شیمیایی عضلات، کبد، قلب و کلیه‌ها را تحت تأثیر قرار می‌دهد و برای رشد و تکامل مغز نیز ضروری‌اند [۱۰]. هورمون‌های تیروئید بر غده هیپوفیز اثر می‌گذارد و موجب ترشح هورمون کنترل‌کننده تیرئید[1] می‌شوند که به نوبه خود موجب تحریک تیروئید می‌شود [۱۱]. وقتی میزان هورمون‌های تیروئید در خون کاهش می‌یابد، هیپوفیز با ترشح بیشتر هورمون کنترل‌کننده تیروئید، تیروئید را فعال‌تر می‌کند [۱۲]. معمولاً تست هورمون کنترل‌کننده تیروئید اصلی‌ترین و مهم‌ترین ارزیابی است، اما همراه با آن ممکن است هورمون تری یدوتیرونین[2] و هورمون تیروکسین[3]، کل مقادیر متصل به پروتئین و همچنین مقادیر آزاد آن (تری یدوتیرونین آزاد[4] و تیروکسین آزاد[5]) مورد بررسی قرار گیرند [۱۳].

به‌دلیل وجود حجم عظیمی از داده‌ها در پایگاه‌های داده پزشکی کشور در مورد افراد مبتلا به بیماری تیروئید، امکان استخراج عوامل پیش‌بینی این بیماری توسط متخصصین بااستفاده‌از استخراج دانش از این حجم عظیم داده عملاً امکان‌پذیر نخواهد بود [۱۴]. بنابراین وجود الگوریتم‌های داده‌کاوی و یادگیری ماشین جهت کشف الگوی‌های پنهان در این داده‌ها و تشخیص و پیش‌بینی این بیماری احساس می‌شود. همچنین پژوهش‌های پیشین در این زمینه مورد بررسی قرار داده شده است.

در مقاله عسگری و خسروی دانش [۱۵، ۱۶] با عنوان تشخیص بیماری تیروئید با استفاده از الگوریتم طبقه‌بندی فازی عصبی، نشان داده شد تشخیص بیماری تیروئید می‌تواند بسیار مشکل باشد چون علائم آن به راحتی با شرایط دیگر اشتباه گرفته می‌شود. بنابراین تشخیص بهنگام و صحیح بی‌نظمی‌های تیروئید بر پایه تست‌های آزمایشگاهی و علائم بیماری امری حیاتی است، سیستم‌های خبره یا سیستم‌های مبتنی بر دانش، رایج‌ترین نوع سیستم‌های (سیستم خبره مبتنی بر هوش مصنوعی[6]) هوش مصنوعی در پزشکی هستند.

معززی و همکاران [۱۷]، تحقیقات خود را بر روی ۴۵ مورد بیمار با درمان لووتیروکسین و ۱۴ نفر سالم انجام دادند. در این راستا آن‌ها هورمون‌های برداشت تری یدوتیرونین (T3RU)[7]، تری یدوتیرونین (T3)، تیروکسین (T4) و هورمون کنترل‌کننده تیروئید قبل از تزریق ۲۵۰ میکروگرم هورمون آزادکننده تیروتروپین[8] و ۳۰ دقیقه بعد از تزریق فقط هورمون کنترل‌کننده تیروئید اندازه‌گیری و به فاصله ۷ تا ۱۱ روز بعد در این افراد آزمون تحمل گلوکز انجام شد. سپس در آخر، یافته‌های به‌دست‌آمده از افراد مبتلا به پرکاری تیروئید و افراد سالم را با هم مقایسه کردند.

هورمون کنترل‌کننده تیروئید TSH

در مقاله ساراواسی و همکاران [۱۸]، تشخیص پزشکی می‌تواند به‌عنوان یک مسئله طبقه‌بندی در نظر گرفته شود. براساس مجموعه‌ای از ویژگی‌های ورودی، هدف این است که یک بیمار را به‌عنوان یک اختلال خاص طبقه‌بندی کند یا اینکه به آن نیاز نداشته باشد. مشکلات هورمون تیروئید در حال حاضر شایع‌ترین مشکلات هستند. در این مقاله یک رویکرد شبکه عصبی مصنوعی بااستفاده‌از یک الگوریتم تکثیر برگشتی برای تشخیص مشکلات تیروئید توسعه داده شده است. تعدادی از عوامل به‌عنوان ورودی هستند و خروجی تولید می‌کنند که نتیجه آن این است که آیا یک فرد مشکل دارد یا سالم است. مشخص شده است که الگوریتم انتشار عقب دارای حساسیت و خاصیت بالا است. همچنین مدل پیشنهادی آن‌ها در نهایت به‌دقت ۹۴ درصد دست یافته است.

1. Thyroid Stimulating Hormone (TSH)
2. Triiodothyronine (T3)
3. Thyroxine (T4)
4. Free Triiodothyronine (FT3)
5. Free Thyroxine (FT4)
6. Artificial intelligence Medical (AIM)
7. Triiodothyronine Resin Uptake (T3RU)
8. Thyrotropin-Releasing Hormone (TRH)

در مقاله پررانا و همکاران [۱۹]، همه موجودات زنده به یک یا چند راه دیگر در معرض اختلال در سلامتی هستند. بیماری تیروئید یک علت اصلی نگرانی در تشخیص پزشکی و پیش‌بینی است، شروع آن در پژوهش‌های پزشکی یک پیشنهاد دشوار است. پیش‌بینی دقیق بیماری با درک علائم و گزارش‌های آسیب‌شناسی، با نگاه کردن به الگوهای تولید و الگوهای گزارش تولیدشده ممکن است. بااستفاده‌از الگوریتم‌های یادگیری ماشین، سیستم پیش‌بینی کامپیوتری را می‌توان بااستفاده‌از شبکه‌های عصبی مصنوعی[۹] توسعه داد. نمونه‌هایی از تیروئید را پس از تمیز کردن و پیش پردازش جمع‌آوری و آن‌ها را به هیپو تیروئید، طبیعی تبدیل کرده‌اند. مجموعه داده‌ها برای آموزش و تست به نسبت ۸۰ به ۲۰ تقسیم شده است. شبکه‌های عصبی مصنوعی اعمال‌شده به این مجموعه در جعبه ابزار شبکه‌های عصبی مصنوعی از نرم‌افزار متلب[۱۰] (زبان برنامه‌نویسی متلب) اجرا شد که با دقت ۸۵ درصد در طی مرحله آموزش مدل انجام شد. این مدل با کلاس‌های مختلف ورودی مورد آزمایش قرار گرفت و درنتیجه ۸۲ درصد دقت داشت.

در مقاله ژانگ و همکاران [۲۰]، بااستفاده‌از شبکه‌های عصبی کانولوشنال عمیق، بااستفاده‌از مقایسه چند طبقه‌بندی با سونوگرافی و توموگرافی کامپیوتری بیماری تیروئید را تشخیص دادند. در این مقاله، از تکنیک‌های مبتنی بر یادگیری عمیق برای تشخیص خودکار بیماری‌های تیروئید استفاده شده است که می‌تواند به پزشکان در تصمیم‌گیری تشخیصی کمک کند. این مطالعه پژوهشی اولین مطالعه در نوع خود است که دو روش تصویر پزشکی قبل از عمل را برای انواع بیماری‌های تیروئیدی چند طبقه‌بندی (یعنی طبیعی، تیروئیدیت، کیستیک، گواتر چند ندولار، آدنوم و سرطان) اتخاذ می‌کند. در این مقاله، بااستفاده‌از معماری شبکه عصبی کانولوشنال[۱۱]، یک مدل تشخیصی بیماری تیروئید را برای تمایز بین انواع بیماری ایجاد می‌کند. مدل موردنظر، دقت بی‌سابقه‌ای را برای مجموعه اسکن‌های سی‌تی‌اسکن به دست می‌آورد. مدل پیشنهادی توانست با دقت ۰/۹۴۲ افراد بیمار را شناسایی کند.

در تمامی پژوهش‌های پیش‌گفت، مدل‌های ارائه‌شده همگی مدل‌های با نظارت بوده‌اند. حال در این پژوهش، سعی شده است بااستفاده‌از روشی یادگیر و بدون نظارت، مبتنی بر تکنیک شبکه عصبی مصنوعی به پیش‌بینی افراد مبتلا به بیماری تیروئید پرداخته شود. در ادامه مدل پیشنهادی و نتایج حاصله تشریح می‌شود.

مواد و روش‌ها

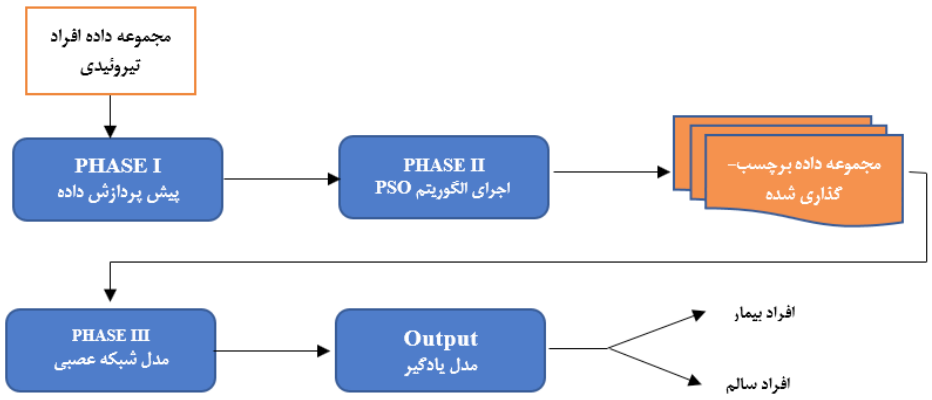
پژوهش حاضر از نوع کاربردی‌پیمایشی بوده و در سال ۱۴۰۱ انجام شده است. مجموعه داده مورداستفاده در این پژوهش، مربوط به بررسی و جمع‌آوری میدانی اطلاعات و نمونه‌گیری هدفمند بیش از ۴۰۰ پرونده بیماران مراجعه‌کننده به بیمارستان امام رضا(ع) شهرستان لارستان از سال ۱۴۰۰ تا ۱۴۰۱ می‌باشد. رویکرد پیشنهادی در این پژوهش، یک رویکرد یادگیری بدون نظارت جهت تشخیص تیروئید در افراد مبتلا به این بیماری می‌باشد. الگوریتم مورد نظر، یک الگوریتم ترکیبی تکاملی حاصل از آمیختگی الگوریتم بهینه‌سازی ازدحام ذرات و شبکه عصبی مصنوعی می‌باشد.

مطابق تصویر شماره ۱، مدل پیشنهادی شامل ۳ فاز اصلی می‌باشد. ابتدا مجموعه داده بیماران تیروئیدی وارد فاز اول جهت پیش پردازش می‌شود. در این فاز مجموعه داده ورودی نرمال‌سازی می‌شوند و رکوردهای ناقص از مجموعه داده حذف می‌شوند. سپس مجموعه داده نرمال‌شده وارد فاز دوم جهت عملیات خوشه‌بندی و برچسب‌گذاری داده‌ها می‌شود. در این فاز ابتدا جهت ایجاد جمعیت ذرات[۱۲] در الگوریتم بهینه‌سازی ازدحام ذرات[۱۳] از الگوریتم کی-میانگین[۱۴] استفاده شد و خروجی آن مراکز ثقل حاصل از خوشه‌بندی با این الگوریتم می‌باشد. سپس جمعیت ذرات وارد الگوریتم بهینه‌سازی ازدحام ذرات جهت یافتن خوشه بهینه می‌شود. بعد از انجام عملیات بهینه‌سازی توسط الگوریتم بهینه‌سازی ازدحام ذرات خوشه‌های بهینه حاوی ۲ خوشه افراد سالم و بیمار (مجموعه داده برچسب‌گذاری‌شده) به‌عنوان خروجی وارد فاز بعد می‌شود. در فاز سوم، مجموعه داده ورودی برچسب‌گذاری‌شده وارد مدل شبکه عصبی طراحی‌شده به منظور فرآیند یادگیری مدل هوشمند می‌شود. شبکه عصبی مورداستفاده در اینجا شبکه عصبی پیش‌خور با یک لایه ورودی، ۲۰ لایه پنهان و ۱ لایه خروجی می‌باشد. بعد از فرآیند آموزش مدل هوشمند موردنظر، درنهایت این مدل به‌عنوان خروجی می‌تواند به‌راحتی افراد دارای بیماری تیروئید را از افراد سالم تشخیص دهد.

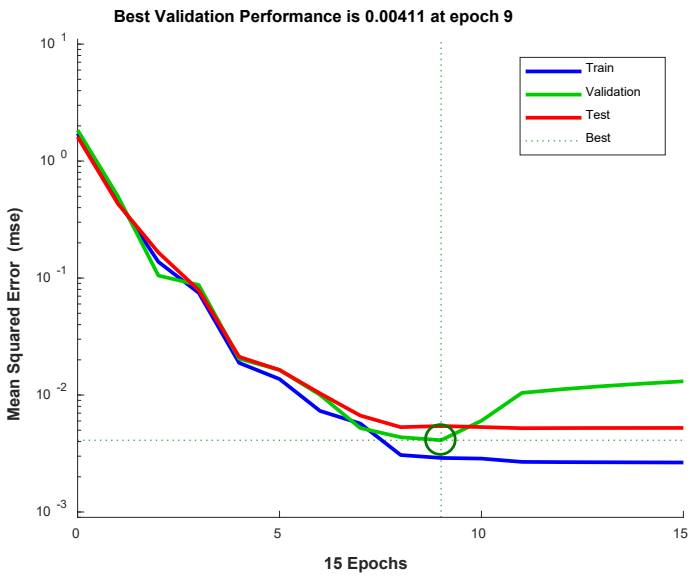
مجموعه داده مورداستفاده

در این پژوهش، سعی شده است بااستفاده‌از یک مدل ترکیبی تکاملی یادگیر، افراد دارای بیماری تیروئید از افراد سالم تشخیص داده شود. مجموعه داده مورداستفاده در این پژوهش، از بیماران مراجعه‌کننده به بیمارستان امام رضا (ع) شهرستان لارستان گرفته شده است. در این مجموعه داده مطابق جدول شماره ۱

9. Artificial Neural Network (ANN)
10. MATLAB
11. Convolutional Neural Network (CNN)
12. particle
13. Particle Swarm Optimization (PSO)
14. K-means



تصویر ۱. فلوچارت مدل پیشنهادی

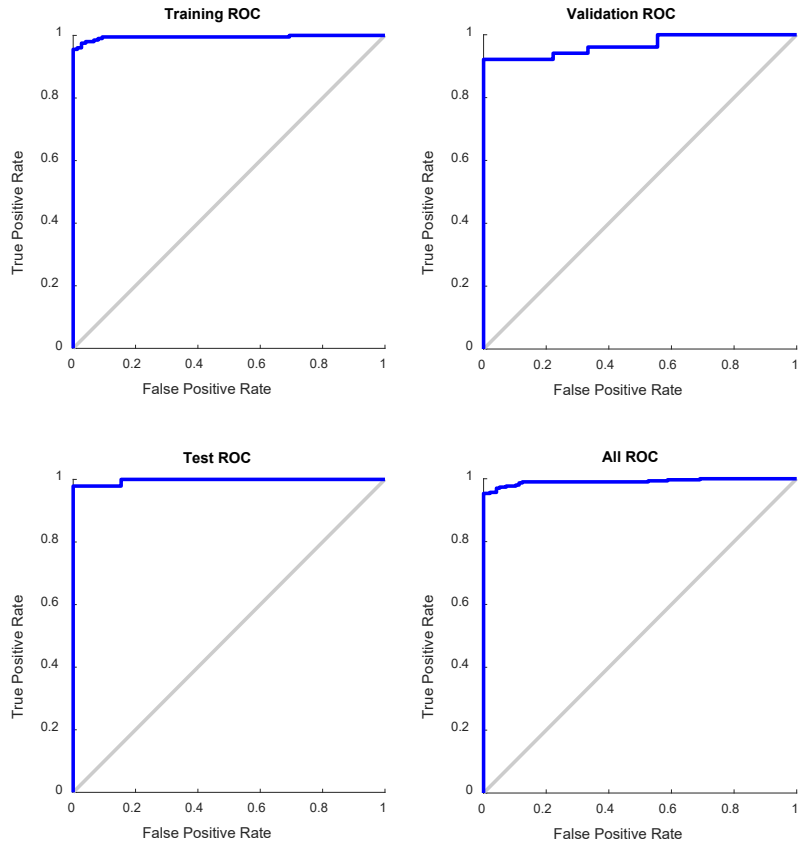


تصویر ۲. مجموع مربعات خطا در مدل پیشنهادی با ۳۰ لایه مخفی

پارامترهای (تیروتروپین سرم، تری یدوتیرونین سرم، تیروکسین سرم و تیروکسین آزاد) در بیش از ۴۰۰ بیمار مورد بررسی قرار گرفته است. مجموعه داده موردنظر شامل ۴ ویژگی اثرگذار در بیماری تیروئید (تیروتروپین سرم، تری یدوتیرونین سرم، تیروکسین سرم و تیروکسین آزاد) می‌باشد که در جدول شماره ۱ محدوده نرمال هریک از این شاخص‌ها ارائه شده است.

جدول ۱. ویژگی‌های اندازه‌گیری‌شده در افراد مراجعه‌کننده

آزمایش	شاخص	محدوده طبیعی
تیروتروپین سرم	TSH	۰/۳۵-۴/۹۴ میکرو واحد بر میلی‌لیتر
تیروکسین آزاد	FT4	۰/۸-۲ نانوگرم بر دسی‌لیتر
تیروکسین سرم	T4	۴/۸۷-۱۱/۷۲ میکروگرم بر دسی‌لیتر
تری‌یدوتیرونین سرم	T3	۰/۵۸-۱/۵۹ نانوگرم بر میلی‌لیتر



تصویر ۳. منحنی راک مربوط به یادگیری سیستم پیشنهادی

ویژگی‌های موردنظر با مشورت پزشک متخصص در نظر گرفته شده است. در این مجموعه داده، ۳۰۰ مورد بیمار تیروئیدی و ۱۰۰ مورد سالم می‌باشند.

یافته‌ها

در این بخش معیارهای مجموع مربعات خطا، منحنی راک[15]، منحنی رگرسیون و ماتریس درهم ریختگی، در مدل پیشنهادی تجزیه‌وتحلیل شدند. برای انجام این‌کار، مدل پیشنهادی بر روی مجموعه‌داده با ۴۰۰ نمونه مورد آموزش قرار گرفته است. در مجموعه داده مورد نظر ۳۰۰ نمونه افراد بیمار و ۱۰۰ نمونه افراد سالم وجود دارد. در ادامه نتایج اجرای مدل پیشنهادی بر روی این مجموعه داده تشریح شده است.

تصویر شماره ۲ مجموع مربعات خطا در ۳ حالت آموزش اعتبارسنجی و تست مدل پیشنهادی را نشان می‌دهد. همان‌طورکه مشاهده می‌کنید، در تکرار ۹ به مقدار ۰/۰۰۴۱۱ رسیده که نشان از کارایی بالای مدل پیشنهادی دارد.

در تصویر شماره ۳ منحنی راک به تصویر کشیده شده است. منحنی راک معیاری است که برای بررسی کیفیت دسته‌بندی‌کننده‌ها مورد استفاده قرار می‌گیرد. برای هر کلاس دسته‌بندی‌کننده، مقادیر آستانه‌ای بین صفر و ۱ برای خروجی‌ها به کار می‌برد. برای هر حدآستانه نیز ۲ مقدار نرخ مثبت واقعی[16] و نرخ مثبت کاذب[17] محاسبه می‌شود. همان‌طورکه در تصویر شماره ۳ مشاهده می‌کنید در مدل پیشنهادی همواره نرخ مثبت واقعی بیشتر از نرخ مثبت کاذب است و مقدار ثبت‌شده نسبت واقعی بر کاذب، ۰/۹۸ می‌باشد

ماتریس اغتشاش مربوط به یادگیری سیستم پیشنهادی مورد بررسی قرار گرفته است. همان‌طورکه در تصویر شماره ۴ مشاهده می‌شود، در ماتریس اغتشاش سطرها نشانگر کلاس پیش‌بینی‌شده (کلاس خروجی) و ستون‌ها نشان‌دهنده کلاس صحیح (کلاس هدف) هستند. سلول‌های مورب نشان می‌دهد چند مقدار (و چه درصدی) از مشاهدات نمونه شبکه آموزش‌دیده را به‌درستی تخمین زده است. همچنین، ۲ سلول مورب اول، تعداد و درصد طبقه‌بندی‌های صحیح را توسط شبکه آموزش‌دیده نشان می‌دهند.

15. (ROC)

16. (TPR)

17. (FPR)

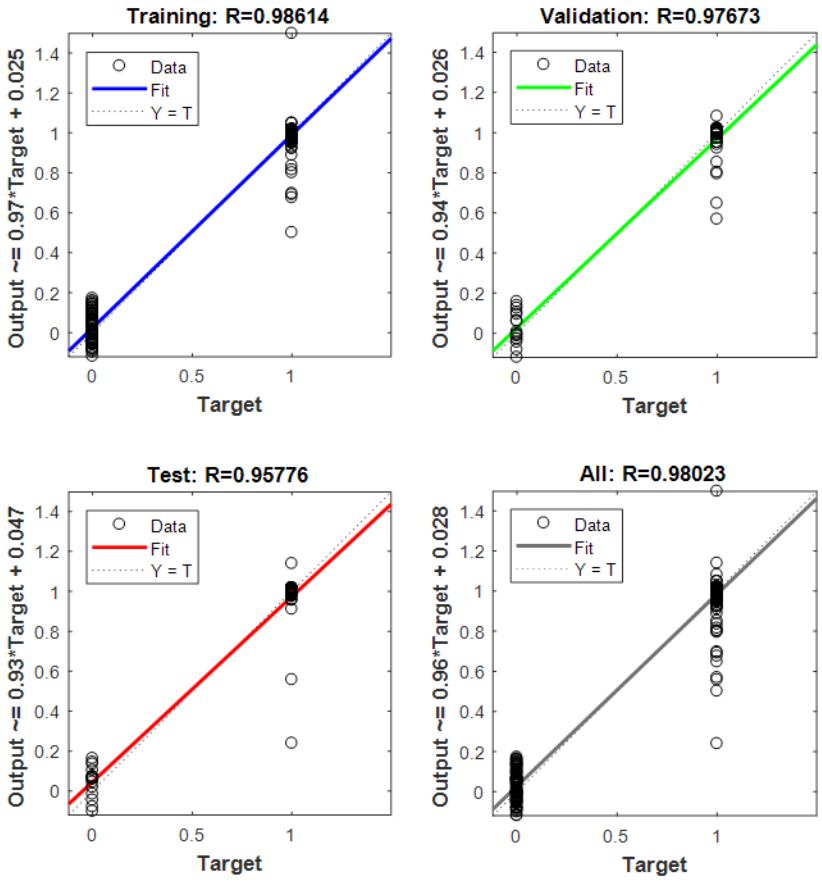
تصویر ۴. ماتریس اغتشاش در سیستم پیشنهادی با ۳۰ لایه مخفی

مطابق با نتایج، ۹۴ نمونه به‌درستی به‌عنوان افراد سالم طبقه‌بندی شده است. این مطلب مطابق با ۲۳/۷ درصد از کل نمونه‌ها است. به‌طور مشابه، ۲۸۷ مورد به‌درستی به‌عنوان افراد بیمار طبقه‌بندی شده‌اند. این مطلب، مطابق با ۷۲/۳ درصد از کل نمونه‌های آزمایشی است. ۱۳ نمونه از افراد سالم به اشتباه بیمار، دسته‌بندی شده‌اند و مطابق با ۳/۳ درصد از کل ۴۰۰ نمونه موجود است. به‌طور مشابه ۳ مورد از افراد بیمار در کلاس سالم قرار گرفته‌اند که مطابق با ۰/۸ درصد از کل نمونه‌ها می‌باشد. مطابق با ماتریس اغتشاش تولیدشده، در مدل پیشنهادی، از ۱۰۷ پیش‌بینی افراد سالم، ۸۷/۹ درصد به‌درستی و ۱۲/۱ درصد به اشتباه کلاس‌بندی شده‌اند. از ۲۹۰ پیش‌بینی افراد بیمار، ۹۹ درصد به‌درستی و درصد به اشتباه کلاس‌بندی شده‌اند. از ۹۷ موارد افراد سالم، ۹۶/۹ درصد به‌خوبی به‌عنوان افراد سالم و ۳/۱ درصد به‌عنوان افراد بیمار پیش‌بینی شده است. از ۳۰۰ مورد افراد بیمار، ۹۵/۷ درصد به‌درستی به‌عنوان بیمار و ۴/۳ درصد به‌عنوان افراد سالم طبقه‌بندی شده‌اند. به‌طورکلی، ۹۶ درصد از پیش‌بینی‌ها درست هستند و ۴ درصد از طبقه‌بندی‌ها اشتباه هستند.

تصویر شماره ۵، منحنی رگرسیون آموزش سیستم پیشنهادی را نشان می‌دهد. همان‌طورکه در شکل مشاهده می‌کنید، ضریب رگرسیون در ۳ حالت آموزش، اعتبارسنجی و تست به‌ترتیب

جدول ۲. مقایسه کارایی روش پیشنهادی در پیش‌بینی بیماری تیرئید با روش‌های بررسی‌شده در پیشینه

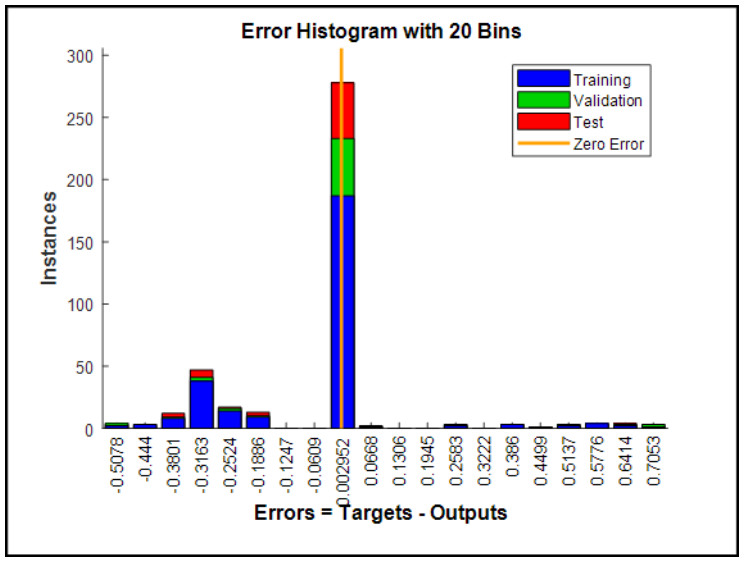
رفرنس	روش	دقت (درصد)
[۱۶]	طبقه‌بندی فازی عصبی	۹۲
[۱۸]	ترکیب الگوریتم تکثیر برگشتی و شبکه عصبی	۹۴
[۱۹]	شبکه‌های عصبی مصنوعی	۸۲
[۲۰]	شبکه عصبی کانولوشنال	۹۴/۴
مدل پیشنهادی	ترکیب الگوریتم بهینه‌سازی ازدحام ذرات و شبکه عصبی	۹۶



تصویر ۵. منحنی رگرسیون آموزش سیستم پیشنهادی

دارای مقادیر (۰/۹۸، ۰/۹۷ و ۰/۹۵) می‌باشد. همچنین ضریب رگرسیون در کل در آموزش سیستم دارای مقدار ۰/۹۸۰۲۳ است. مقدار ضریب رگرسیون هر اندازه به عدد ۱ نزدیکتر باشد، نشان‌دهنده بهینه بودن سیستم پیشنهادی می‌باشد بنابراین به‌دلیل نزدیک بودن مقدار ضریب رگرسیون سیستم پیشنهادی به عدد ۱، پس عملکرد سیستم پیشنهادی مطلوب می‌باشد

تصویر شماره ۶ هیستوگرام نرخ خطا در یادگیری سیستم پیشنهادی را به‌تصویر کشیده است. همان‌طورکه مشاهده می‌کنید نرخ خطا در سیستم پیشنهادی بسیار پایین و نزدیک



تصویر ۶. هیستوگرام نرخ خطا در یادگیری سیستم

به صفر می‌باشد و در اکثر نمونه‌ها نرخ خطا نزدیک به صفر و اختلاف مقدار واقعی با مقدار پیش‌بینی‌شده مقدار ۰/۰۰۲۹ است. مقدار پایین نرخ خطا نشان‌دهنده عملکرد بهینه سیستم پیشنهادی می‌باشد

مقایسه نتایج روش پیشنهادی با روش‌های پیشین مطابق با معیار دقت ارائه شده است. مطابق با جدول شماره ۲، مدل پیشنهادی قادر به پیش‌بینی بیماری تیروئید با دقت ۹۶ درصد می‌باشد که این میزان دقت برتری روش پیشنهادی نسبت به روش‌های پیشین می‌باشد

بحث

در این تحقیق، رویکردی هوشمند و بدون نظارت، مبتنی بر الگوریتم بهینه‌سازی ازدحام ذرات و شبکه‌عصبی مصنوعی ارائه شده است. مدل پیشنهادی به دنبال بهبود فرآیند پیش‌بینی بیماری تیروئید در افراد می‌باشد. بخش اصلی مدل ترکیبی پیشنهادی در این پژوهش، شبکه‌عصبی‌مصنوعی می‌باشد. شبکه‌عصبی موردنظر بااستفاده‌از مجموعه داده‌های برچسب‌گذاری‌شده‌ای که در اختیار دارد، سعی می‌کند الگوهای پنهان‌شده افراد بیمار از افراد سالم را پیدا کند و به این ترتیب افراد دارای بیماری تیروئید را در این مجموعه داده استخراج می‌کند. ازآنجاکه نحوه عملکرد شبکه عصبی به چگونگی تعریف ساختار آن بستگی دارد، در این پژوهش در کنار شبکه عصبی پیشنهادشده، از الگوریتم بهینه‌سازی ازدحام ذرات استفاده شده است که وظیفه آن نحوه برچسب‌گذاری بهینه مجموعه داده ورودی می‌باشد تا بدین نحو بستر یادگیری بدون نظارت در ساختار شبکه‌عصبی موردنظر را به وجود آورد که برتری روش پیشنهادی نسبت به روش‌های پیشین است. چون روش‌های پیشین همگی از یک مدل یادگیر با نظارت جهت انجام عملیات تشخیص استفاده کرده‌اند.

مدل پیشنهادی در محیط متلب پیاده‌سازی و عملکرد آن موردارزیابی قرار گرفته است. بررسی نتایج نشان می‌دهد مقادیر ضریب رگرسیون (R) مدل پیشنهادی در ۳ حالت آموزش، اعتبارسنجی و تست به‌ترتیب دارای مقادیر (۰/۹۸، ۰/۹۷ و ۰/۹۵)، منحنی راک برابر با ۰/۹۸ و میزان خطا ۰/۰۰۴ بوده است. همچنین روش پیشنهادی با روش‌های بررسی‌شده در پیشنه پژوهش براساس معیار دقت مورد بررسی و تجزیه‌وتحلیل قرار گرفته است. نتایج به‌دست‌آمده حاکی از این است که مدل پیشنهادی قادر است با دقت ۹۶ درصد بیماری تیروئید را در بیماران به‌درستی پیش‌بینی کند. این در حالیست که مدل‌های پیشین نظیر مدل پیشنهادی پررانا و همکاران که یک روش مبتنی بر شبکه عصبی مصنوعی با نظارت است، درنهایت به‌دقت ۸۲ درصد قادر به پیش‌بینی این بیماری می‌باشد.

همچنین مدل پیشنهادی ساراواسی و همکاران نیز توانسته با دقت ۹۴ درصد بیماران تیروئیدی را تشخیص دهد؛ مدل پیشنهادی ژانگ و همکاران نیز که براساس روش شبکه عصبی کانولوشنال بوده است، درنهایت با دقت ۹۴ درصد قادر به تشخیص بیماران تیروئید می‌باشد. درنتیجه، مدل پیشنهادی در این پژوهش با ساختار بدون نظارت توانسته به‌دقت ۹۶ درصد برسد، درحالی‌که هرکدام از مدل‌های پیشن با یک مدل با نظارت توانسته بودند به چنین دقت‌هایی دست پیدا کنند. بنابراین نتایج پیاده‌سازی نشان می‌دهد عملکرد و دقت سیستم پیشنهادی در تشخیص و پیش‌بینی بیماری تیروئید از روش‌های پیشین بهتر است و عملکرده قابل قبولی دارد.

از محدودیت‌های پژوهش استفاده از الگوریتم بهینه‌سازی ازدحام ذرات در کنار شبکه‌عصبی‌مصنوعی شده است. چون امکان گیر کردن در بهینه محلی وجود دارد. احتمال گیر افتادن در تله بهینه محلی کم است، اما پیشنهاد می‌شود در مدل پیشنهادی از الگوریتم‌های تکاملی دیگر نظیر ژنتیک، الگوریتم کرم شب‌تاب، الگوریتم خفاش و غیره به‌جای الگوریتم بهینه‌سازی ازدحام ذرات جهت انجام عملیات خوشه‌بندی بدون نظارت و همچنین برای یادگیری وزن‌های شبکه‌عصبی استفاده شود و نتایج مورد تجزیه‌وتحلیل قرار گیرند.\

نتیجه‌گیری

باتوجه‌به نتایج حاصله و ارزیابی آن‌ها، می‌توان به‌طور کلی نتیجه گرفت که مدل پیشنهادی می‌تواند با دقت قابل قبولی، بیماری تیروئید را در افراد پیش‌بینی کند. بدین ترتیب میزان اشتباه در تشخیص و پیش‌بینی این بیماری را تا حد قابل توجهی کاهش دهد. همچنین از این مدل می‌توان به‌عنوان یک ابزار مفید در پیش‌بینی بیماری تیروئید در مراکز مختلف درمانیا ستفاده کرد.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

این مطالعه دارای تأییدیه اخلاقی به شماره IR.IAUESTAHBAN.REC.1401.023 از دانشگاه آزاد اسلامی واحد استهبان است.

حامی مالی

این پژوهش هیچ‌گونه کمک مالی از سازمانی‌های دولتی، خصوصی و غیرانتفاعی دریافت نکرده است.

تعارض منافع

بنابر اظهار نویسنده این مقاله تعارض منافع ندارد.

References

[1] Rajam K, Jemina Priyadarsini R. A survey on diagnosis of thyroid disease using data mining techniques. Int J Comput Sci Mob Comput. 2016; 51(5):354-8. [Link]

[2] Hosseinpoor M, Parvin H, Nejatian S, Rezaee V, Bagherifard K. [Presenting a meta-heuristic algorithm to detect regulatory elements in the genome of breast cancer patients (Persian)]. J Adv Biomed Sci. 2020; 10(1):2070-80. [Link]

[3] Dabiri MR, Farshidfar A, Najaflo Sh, Saadatdoost R. [Diagnosing thyroid disease using meta-heuristic hybrid algorithms (Persian)]. Paper presented at: Third International Conference on Electrical, Computer and Mechanical Engineering. 20 October 2020; Tehran, Iran. [Link]

[4] Hosseinpoor M, Parvin H, Nejatian S, Rezaei V. [Detection and extraction of potential promoter/enhancer interactions in genome of cancer patients using an evolutionary multi-objective algorithm (Persian)]. J Health Biomed Inform. 2018; 5(2):304-13. [Link]

[5] Jalili S, Faraji Daneshgar F. [Diagnosis of thyroid diseases using RFC fuzzy classification method (Persian)]. Paper presented at: The 14th Iran Medical Engineering Conference. 13 February 2008; Tehran, Iran. [Link]

[6] Hosseinpoor M. [Predicting gestational diabetes using an intelligent algorithm based on artificial neural network (Persian)]. J Mod Med Inf Sci. 2022; 8(2):126-39. [Link]

[7] Zabbah I, Yasrebi Naeini SE, Ramazanpoor Z, Sahragard K. [The diagnosis of thyroid diseases using combinati on of neural networks through hierarchical method (Persian)]. J Health Biomed Inform. 2017; 4(1):21-31. [Link]

[8] Aversano L, Bernardi ML, Cimitile M, Iammarino M, Macchia PE, Nettore IC, et al. Thyroid disease treatment prediction with machine learning approaches. Procedia Comput Sci. 2021; 192:1031-40. [DOI:10.1016/j.procs.2021.08.106]

[9] Upadhayay A, Shukla S, Kumar S. Empirical comparison by data mining Classification algorithms (C 4.5 & C 5.0) for thyroid cancer data set. Int J Comput Sci Commun Netw. 2013;3(1):64. [Link]

[10] Gharehchopogh FS, Molany M, Mokri FD. Using artificial neural network in diagnosis of thyroid diasise: A case study. International Journal on Computational Sciences & Applications. 2013; 3(4):49-61. [Link]

[11] Dhaygude P, Handore SM. A review of thyroid disorder detection using medical images. Int J Recent Innov Trends Comput Commun. 2014; 2(12):4194-7. [Link]

[12] Dogantekin E, Dogantekin A, Avci D. An expert system based on Generalized Discriminant Analysis and Wavelet Support Vector Machine for diagnosis of thyroid diseases. Expert Syst Appl. 2011; 38(1):146-50. [DOI:10.1016/j.eswa.2010.06.029]

[13] Zhang X, Lee VCS, Rong J, Liu F, Kong H. Multi-channel convolutional neural network architectures for thyroid cancer detection. PLoS One. 2022; 17(1):e0262128. [DOI:10.1371/journal.pone.0262128] [PMID]

[14] Abdolali F, Kapur J, Jaremko JL, Noga M, Hareendranathan AR, Punithakumar K. Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks. Comput Biol Med. 2020; 122:103871. [DOI:10.1016/j.compbiomed.2020.103871] [PMID]

[15] Asgari N, Khosravi Danesh F. [Diagnosis of thyroid disease using neural fuzzy classification algorithm (Persian)]. Paper presented at: Second National Conference of Computer Engineering and Information Technology Management. 14 Feruary, 2014; Tehran, Bo Ali Research Group, Iran. [Link]

[16] Yazdan SA, Ahmad R, Iqbal N, Rizwan A, Khan AN, Kim DH. An efficient multi-scale convolutional neural network based multi-class brain MRI classification for SaMD. Tomography. 2022; 8(4):1905-27. [DOI:10.3390/tomography8040161] [PMID]

[17] Moazezi Z, Hedayati M, shirkhani Z, Azizi F. [Glucose intolerance in subclinical hyperthyroid patients (Persian)]. Iran J Endocrinol Metab. 2012; 14(2):127-34. [Link]

[18] Saraswathi V, Santhakumaran A. Towards artificial neural network model to diagnose thyroid problems. Glob J Comput Sci Technol. 2011; 11(5):53-5. [Link]

[19] Prerana PS, Taneja K. Predictive data mining for diagnosis of thyroid disease using neural network. Int J Res Manag Sci Technol. 2015; 3(2):75-80. [Link]

[20] Zhang X, Lee VC, Rong J, Lee JC, Liu F. Deep convolutional neural networks in thyroid disease detection: A multi-classification comparison by ultrasonography and computed tomography. Comput Methods Programs Biomed. 2022; 220:106823. [PMID]