Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Exploring the Challenges of Diagnosing Thyroid Disease with Imbalanced Data and Machine Learning: A Systematic Literature Review

Dhekre Saber Saleh*, ***Mohd Shahizan Othman***

Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia
*Corresponding Author.

Received 07/02/2023, Revised 16/06/2023, Accepted 18/06/2023, Published Online First 20/08/2023

Abstract

Thyroid disease is a common disease affecting millions worldwide. Early diagnosis and treatment of thyroid disease can help prevent more serious complications and improve long-term health outcomes. However, thyroid disease diagnosis can be challenging due to its variable symptoms and limited diagnostic tests. By processing enormous amounts of data and seeing trends that may not be immediately evident to human doctors, Machine Learning (ML) algorithms may be capable of increasing the accuracy with which thyroid disease is diagnosed. This study seeks to discover the most recent ML-based and data-driven developments and strategies for diagnosing thyroid disease while considering the challenges associated with imbalanced data in thyroid disease predictions. A systematic literature review (SLR) strategy is used in this study to give a comprehensive overview of the existing literature on forecasting data on thyroid disease diagnosed using ML. This study includes 168 articles published between 2013 and 2022, gathered from high-quality journals and applied meta-analysis. The thyroid disease diagnoses (TDD) category, techniques, applications, and solutions were among the many elements considered and researched when reviewing the 41 articles of cited literature used in this research. According to our SLR, the current technique's actual application and efficacy are constrained by several outstanding issues associated with imbalance. In TDD, the technique of ML increases data-driven decision-making. In the Meta-analysis, 168 documents have been processed, and 41 documents on TDD are included for observation analysis. The limits of ML that are discussed in the discussion sections may guide the direction of future research. Regardless, this study predicts that ML-based thyroid disease detection with imbalanced data and other novel approaches may reveal numerous unrealised possibilities in the future

Keywords: Classification, Deep learning, Imbalanced data, Machine learning, Thyroid disease.

Introduction

The thyroid gland, which produces hormones that regulate metabolism, is affected by thyroid disease. Nevertheless, it is the most common type of cancer that affects the endocrine system[1]. In recent decades, there has been a rise in the prevalence of thyroid cancer, particularly in countries with advanced diagnostic techniques and reasonable access to medical care. The dominant form of thyroid cancer is differentiated thyroid cancer (DTC), further categorized into papillary and follicular thyroid carcinoma[2]. Whilst the overall prognosis for differentiated thyroid cancer (DTC) is typically favorable, managing this condition may lead to prolonged morbidity due to the elevated risk of recurrence and potential surgical complications[3]. There are two main categories of thyroid disease: hypothyroidism, characterized by a decrease in thyroid hormones and symptoms such as weight

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

gain, fatigue, and constipation, and hyperthyroidism, characterized by an excess of thyroid hormones and symptoms such as irritability, weight loss, and tremors. A combination of environmental factors, genetic factors, and the interaction between these two can cause thyroid disease. Environmental triggers and genetics are potential causes that can contribute to the development of thyroid disease, according to the Mayo Clinic[4].

Additionally, infertility in women may be caused by thyroid gland diseases such as hypothyroidism, hyperthyroidism, and other thyroid gland disorders[5]. When the thyroid gland does not generate enough hormones, it causes weariness, weight gain, and sadness[6]. Conversely, hyperthyroidism happens when the thyroid gland produces too many hormones, resulting in weight loss, anxiety, and tremors[7]. Both disorders can seriously affect a person's health and quality of life.

Certain medications or issues with the pituitary gland regulating the thyroid can also cause thyroid disease. The high annual mortality rate from thyroid disease highlights the significant impact of thyroid cancer on global health[8]. To facilitate clinical decision-making, developing decision models that account for the different causes of death that may compete with thyroid cancer is essential[9]. At the same time, technology can improve healthcare delivery and strengthen health infrastructure [10]. In addition, early detection and timely treatment of thyroid disease can reduce fatalities[11]. The diagnosis of thyroid disease entails a battery of examinations encompassing blood analyses, imaging procedures, and biopsies. The TSH blood test is a primary screening tool for diagnosing thyroid disorders. Medical imaging modalities, including ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI), can produce high-resolution visual representations of the thyroid gland and its adjacent anatomical structures. The medical procedure known as biopsy involves the extraction of a minute tissue sample for laboratory analysis. This procedure is conducted to verify a diagnosis of thyroid cancer[12]. However, healthcare providers may face challenges in diagnosing thyroid disease due to similar symptoms to other illnesses, variable symptoms, limited access to specialized care, and limited diagnostic tests[13].

With the advancement of machine learning in healthcare, many experts consider Thyroid Disease Diagnosis Based on Machine Learning (TDDBML) a viable option. Machine learning improves the accuracy and efficiency of disease diagnosis, and algorithms can scan large amounts of data and recognize patterns that doctors may overlook. ML algorithms can evaluate electronic health data and patient monitoring devices to identify early indications of disease[14]; algorithms can scan high amounts of data and recognize patterns that doctors may miss—improving patient care and reducing wait times. The prevalence of thyroid disease and its significant impact on public health have led to the exploration of using ML for its diagnosis[15].

Several investigations have suggested using ML algorithms, including Support Vector Classifier (SVC), Artificial Neural Networks (ANN), Naive Bayes, Random Forest, and K-Nearest Neighbors, for diagnosing thyroid diseases using various datasets. For instance, Islam S et al, found that the ANN classifier achieved a 96% accuracy rate in predicting thyroid diseases[16]. In addition, Vairale et al, used classification machine learning technique SVM to predict the level of hypothyroid disease[17] while Shyamala Devi et al, used multiple ML techniques with a 99% accuracy rate for predicting hypothyroid disease[18]. Guleria et al, achieved 100% accuracy in the early prediction of hypothyroidism using ANN[19].

One of the possible limitations of ML and deep learning (DL)-based solutions is that they frequently involve sophisticated algorithms that require a large amount of data to train This can make it hard for doctors to evaluate the algorithm's diagnosis and raises bias and reliability concerns[20]. For instance, DLs have numerous invisible layers, but it is not always easy to tell what role each plays in the model's predictions[21]. Another potential difficulty is that ML algorithms tend to support the majority class in their results. The term majority class refers to a dataset in which one category leads the others in a total value [22]. Therefore, it is essential for researchers and healthcare providers to carefully consider these issues when developing and using ML-based models to predict thyroid disease to ensure that they are unbiased.

Table 1 summarizes the review studies that focus on applying ML techniques in TDD and highlights the limited use of SLR compared to the focus on ML techniques. For instance, the K Lee et al. study presents an SLR with a drawback: the machine learning methods applied vary with the data used for thyroid disease diagnosis[23]. Nevertheless, since most datasets used for thyroid disease

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

diagnosis are imbalanced, evaluating the performance of ML on such data is crucial. The study lacked a precise specification of the periods involved[24]. With the increasing popularity of ML-based diagnosis, applying SLR with meta-analysis is expected to address the gaps in existing studies.

The increasing number of studies in Thyroid Disease Diagnosis Based on Machine Learning (TDDBML) highlights the need for a systematic review of existing knowledge. An SLR was conducted using Scopus and WoS databases, resulting in the analysis of 168 papers and further examination of 41—the metadata analysis aimed to identify leading academic institutions, critical research areas, and high-quality sources. In addition, a comprehensive review of 41 publications was conducted to address the following inquiries: What are the existing DL and ML-based approaches for diagnosing thyroid disease? What are the current techniques for dealing with datasets with an imbalanced class ratio?

The SLR aims to supply a resource for researchers by summarizing the latest methods and developments in the field and identifying gaps in knowledge that may be addressed by creating a more advanced TDDBML model. The structure of the remaining article is a methodology of the systematic literature review, briefly described in Section 2. Section 3 presents the results and analysis, Section 4 summarizes the findings, and Section 5 contains the conclusion.

Table 1. Related research for TDDBML

Paper title	Date Range	Study focus	Algorithm	Imbalance challenge	Meta-analysis	Content analysis	SLR
Machine Learning on thyroid disease: a review [23]	2020 or later	ML algorithms differ in data.	✓				
Application of Data Mining Techniques in Diagnosing Various Thyroid Ailments: A Review [24]	Not specified	data mining methods	✓				
review of Deep Learning Approaches for Thyroid Cancer Diagnosis [25]	2018 or late	Estimating diagnostic accuracy of deep learning	✓				
Our study	2013 to 2022	thyroid disease diagnosis based on machine learning	✓	✓	✓	✓	✓

Materials and Methods

An SLR positions research questions before systematically searching for, selecting, and evaluating studies to see what information they may obtain [26]. This approach is chosen due to its reputation for providing a precise and reliable synthesis of scholarly content and is widely recognized across diverse research fields. The eligible study items for meta-analyses, and in this study, the PRISMA guidelines were followed for conducting and reporting the systematic review. The PRISMA checklist was used to ensure that all relevant information was included in the study, and the flow diagram was utilized to document the study selection process[27].

Identification of the data

A thorough exploration was conducted using Scopus's integrated and WoS databases, which include all major publishers, including Emerald, Taylor & Springer, IEEE, and Willey. Many researchers consider the WoS and Scopus databases reliable for SLR due to the excellent quality of the indexing contents [28]. The search covers 2013 to 2022 and includes all essential papers published during this time, employing keywords such as “thyroid", "machine learning", "imbalance", and "deep learning" to find relevant publications. In addition, Boolean operators and various keywords are used to improve the search.

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Screening initial data and determining eligibility

The Scopus and WoS databases were extensively searched for this systematic literature review using specific keywords and a query that included "thyroid disease" OR "thyroid" AND "machine learning" OR "artificial intelligence" OR "data mining" OR "deep learning" AND "diagnosis" OR "detection" OR "classification" OR "prediction." The initial Scopus search revealed 2,182 articles, whereas the WoS search yielded 486. After applying the year boundaries of 2013-November-2022 and further filtering based on document type, language, subject area, and keyword constraints, the number of papers was decreased to 168. Afterwards, the remaining 168 distinct papers were assessed, and the most pertinent information was extracted using a consistent extraction template. The study excluded papers that were not related to machine learning or were primarily concerned with thyroid disorders.

Furthermore, book chapters, ultrasound imaging, non-human studies, and reviews were not included. Finally, 41 full-text papers met the inclusion criteria in Fig 1 and were included in the review. A flowchart was created to show the study selection process, including the search query and inclusion criteria. Overall, the study selection approach was thorough and rigorous, ensuring that the analysis included the most relevant and latest studies on using machine learning techniques to identify thyroid disorders.

The inclusion and exclusion criteria were constructed using machine learning techniques to ensure that the review was comprehensive and relevant to thyroid disease diagnosis. Articles and conference papers about the TDDBML, original research findings or empirical data about the TDDBML, machine learning and profound learning studies about the TDDBML published between 2013 and 2022 met the inclusion requirements. Papers that have not yet been published in English, machine learning studies unrelated to TDDBML, published literature previous to 2013, non-article or conference paper research, duplicate studies, preliminary data studies, or studies with confusing conclusions were also excluded. These criteria were used to employ machine learning techniques to examine the most relevant and up-to-date papers on the diagnosis of thyroid disease while removing irrelevant studies that did not fulfil the conditions mentioned above. The objective of the literature review was to use machine learning techniques to explore the latest and most pertinent publications concerning the diagnosis of thyroid disease.

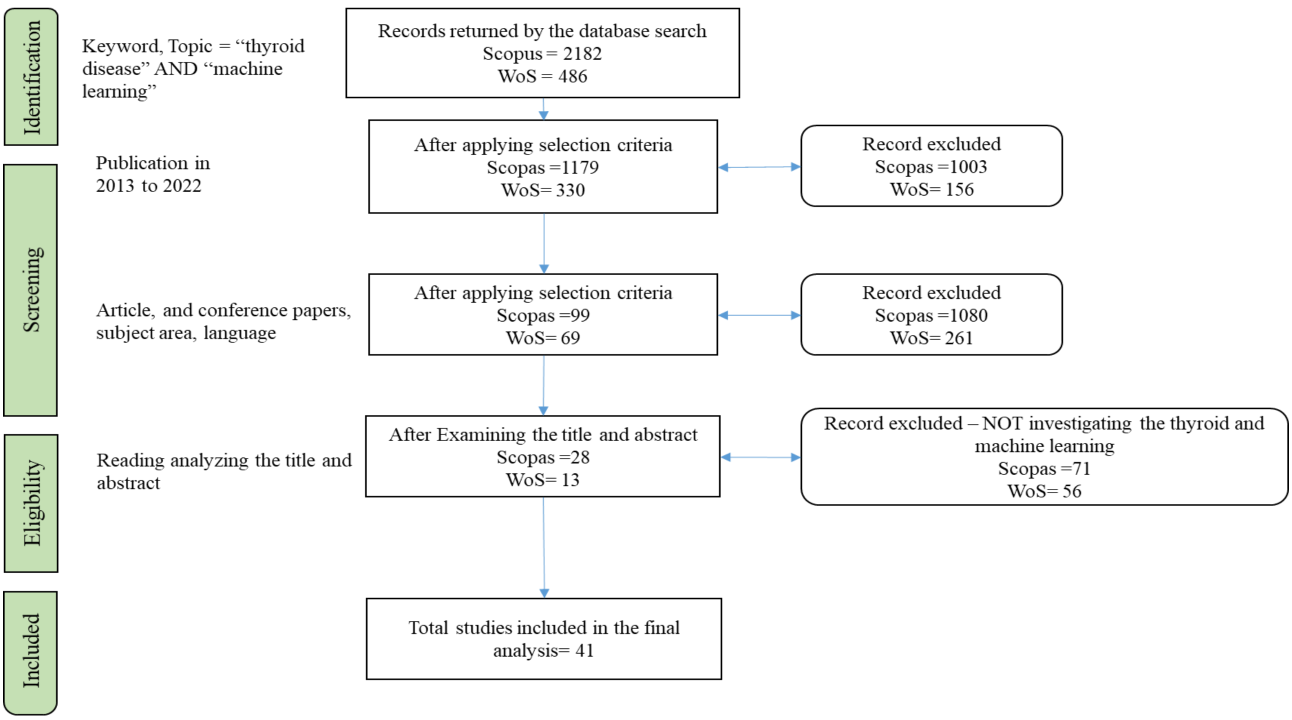


Figure 1. PRISMA approach applied in this research.

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986

Observations and findings

The following section will discuss the findings and insights from evaluating the metadata. These results are based on a meta-study of 168 papers, including studying their corresponding metadata and content.

Metadata analysis

Metadata analysis helps understand scholarly literature by extracting information about the scholarly process authors, articles, journals, and other elements[29]. The metadata analysis was applied to 168 papers. The papers were classified based on various factors, including year of publication, publication type, publisher, country of origin, subject matter, funding source, and academic institution.

Published by year

As shown in Figure 2, 168 papers were reviewed to see how many dealt with thyroid disease prediction using ML algorithms over the past decade. Publishing is constantly expanding, and this increase is expected to intensify significantly in 2020 and 2022. For example, in 2022, around 60 new papers were published; in 2020, just 36 new papers were published.

In addition, it has been evident throughout the period that the significance of the classification problem in the diagnosis of thyroid disease has received much attention. As a result, the number of scholarly works distributed to the public in 2022 is substantially more than in any previous year. On the other hand, one can observe the minuscule number of papers published, particularly from 2013 to 2017, when there were only a handful of papers. Consequently, increasing focus and concern are directed toward diagnosing thyroid disease, including classification issues and other data-driven concerns.

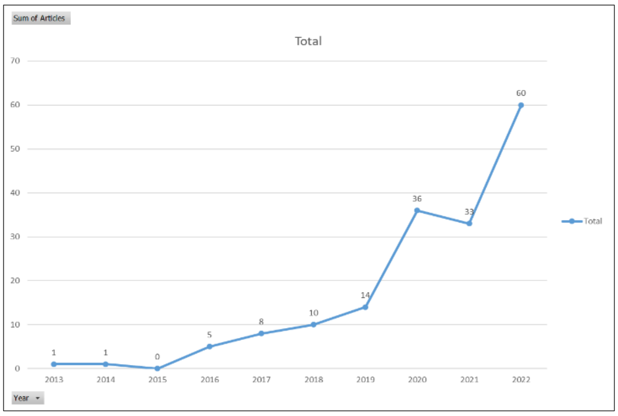
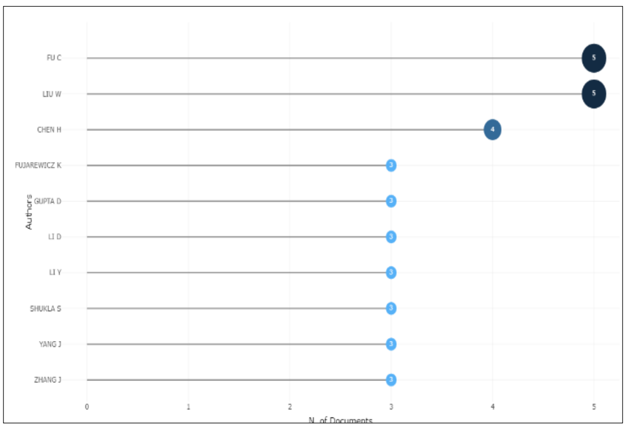


Figure 2. Published papers on thyroid disease by using ML

Most Relevant Authors

According to Figure 3, Fu C and Liu W. have penned the most pertinent papers of the five and are the most impactful authors. Therefore, our team conducted a comparable data examination to monitor the author's production over time. Indeed, findings revealed that Fu C and Liu W jointly produced four 2021 articles that received 10.5 citations.



Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Figure 3. Most relevant authors. The x-axis indicates the number of documents; the y-axis indicates

Most Relevant Sources

As shown in Figure 4, the most pertinent sources had ten documents: advances in intelligent systems and Computers, expert systems with Applications, a total of 6, and The Journal of the study material in Networks and Systems, a total of 5 similarly.

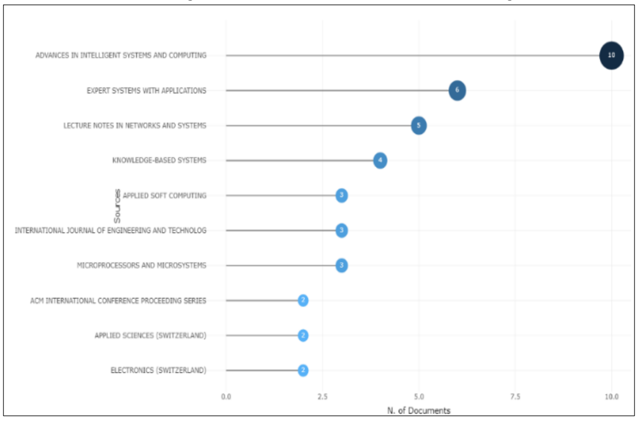


Figure 4. Most relevant sources. The x-axis shows the number of documents; the y-axis indicates

Most frequently words used in the titles and keywords Table

The most common single, double, and triple keywords employed in the paper titles are shown in Table 2. The R-software program is used to identify the most popular keywords. Our main goal was to find and evaluate articles on machine learning, deep learning, imbalance class, and thyroid disease. However, it was surprising to find that the most frequently used keywords in the articles were "thyroid" and "disease", as shown in

Table 3. The writers used the terms "machine learning" 31 times in the keyword field, followed by the terms "thyroid disease" 24 times and "classification" 19 times. Of course, articles often use the exact phrases which are listed below. However, intriguing outcomes were discovered when the examination was limited to the keywords applied by the authors in the articles' keyword sections. A word cloud is a straightforward method for identifying the prevalent themes and key phrases in the referenced articles, allowing for identifying the most general terms in a complex environment. Figure 5 displays word clouds generated by software, where larger and bolder text represents the terms most frequently used, and smaller and less bold text highlights the less commonly used phrases.

Table 2. Most frequently applied words in keyword sections

Unigrams	Frequency	Bigrams	Frequency	Trigrams	Frequency
thyroid	86	thyroid disease	29	thyroid disease prediction	5
learning	43	machine learning	25	Machine learning algorithms	5
classification	42	Deep Learning	9	support vector machine	4
disease	36	feature selection	9	deep neural network	3
machine	32	thyroid cancer	9	machine learning models	3
Data	26	neural network	8	machine learning techniques	3
Based	24	disease diagnosis	6	medical data classification	3

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Unigrams	Frequency	Bigrams	Frequency	Trigrams	Frequency
diagnosis	23	disease prediction	5	thyroid disease classification	3
cancer	18	learning algorithms	5	thyroid disease diagnosis	3
Deep	18	support vector	5	artificial neural network	2

Table 3. Terms repeated in articles

Terms	Frequency
machine learning	31
thyroid disease	24
classification	19
Thyroid	18
feature selection	16
deep learning	15
data mining	14
random forest	10
decision tree	9
thyroid cancer	8

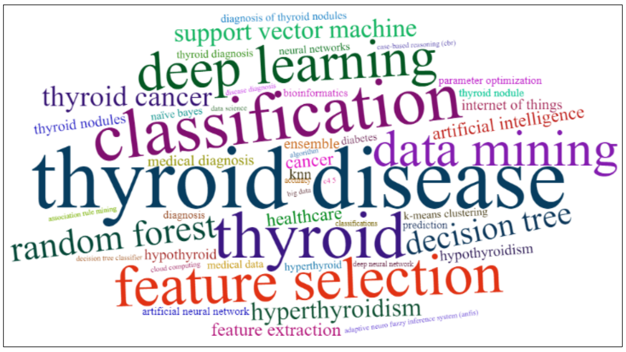


Figure 5. Word cloud for most frequently used keywords in thyroid disease publications.

Trending Topics

The trend topics were generated by introducing only papers published from 2013 to 2022. Graphical parameters, including the author's keywords field, were used, with a minimum word frequency of three and three words considered per year. From Fig 6, the main keywords used each year can be observed. The lines showed when each word was used, and the size of the bubbles indicates how frequently the term appeared. For example, the most frequently used term in 2021 was "machine learning".

Interestingly, the trend in research has evolved over the years. In 2021, the most frequent word was "deep learning," and then the research shifted towards exploring "thyroid" in 2022, followed by "thyroid disease" and "classification" Over the last several years, other words such as "feature selection," "thyroid cancer," and "artificial intelligence" have appeared.

The data reveals that machine learning and deep learning have become increasingly popular, with a specific interest in applying them as a healthcare model for disease diagnosis. Thyroid disease and thyroid cancer are the most researched thyroid-related topics, likely due to their high prevalence worldwide. Understanding and diagnosing thyroid disorders is crucial in the medical field, and these research trends highlight the importance of applying advanced technologies to improve patient care and treatment. Overall, the data provide exciting insights into the evolving research trends in machine learning, deep learning, and thyroid-related topics.

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

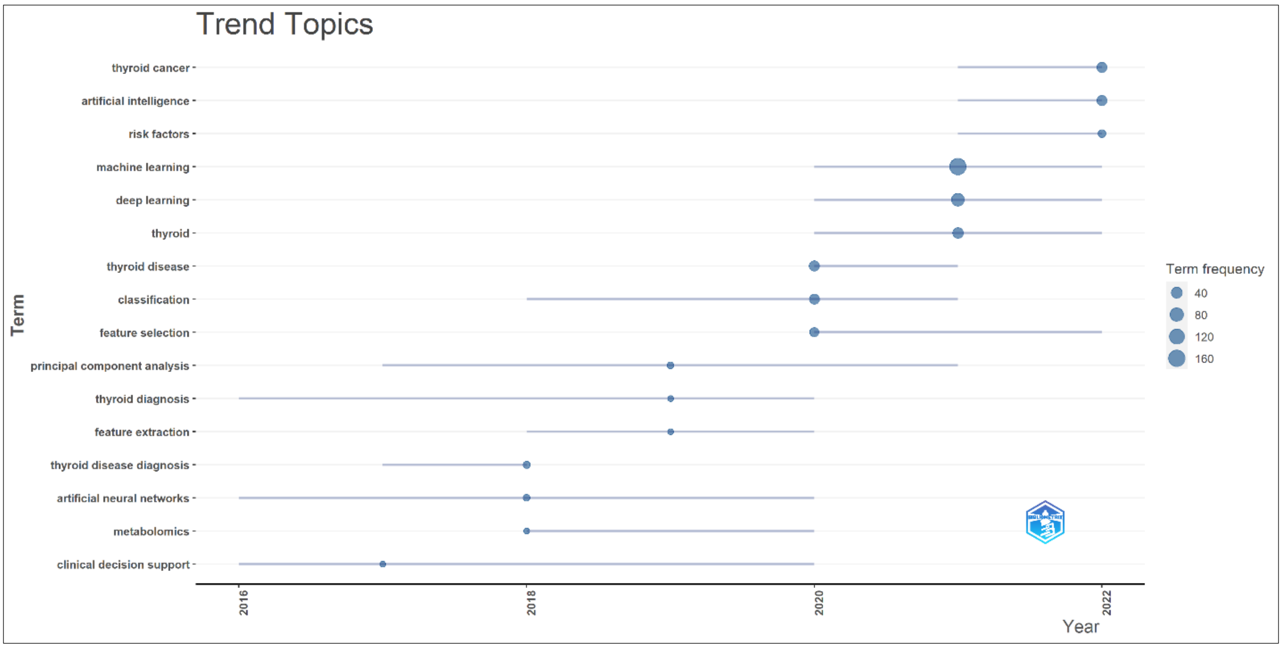


Figure 6. Trending topics extracted from the topic of thyroid disease prediction

Publication by institutions

Fig 7 displays the top 10 academic affiliations worldwide with the most published articles. The list predominantly comprises universities from the United States and China, each having four and three universities, respectively. Other countries such as Saudi Arabia, Pakistan, and Poland are also represented. Zhejiang University in China leads with 18 published articles, followed by Stanford University School of Medicine with 15 articles. This information suggests that Zhejiang University is a leading institution in the field, potentially due to its emphasis on research and development.

The data also provides a snapshot of the current research output from academic affiliations worldwide. It highlights the ongoing efforts of universities to produce high-quality research that can enhance our understanding of various fields and contribute to developing new knowledge and innovations. Understanding the institutions producing the most research in a particular field can be valuable for researchers to make informed decisions and arrange research efforts.

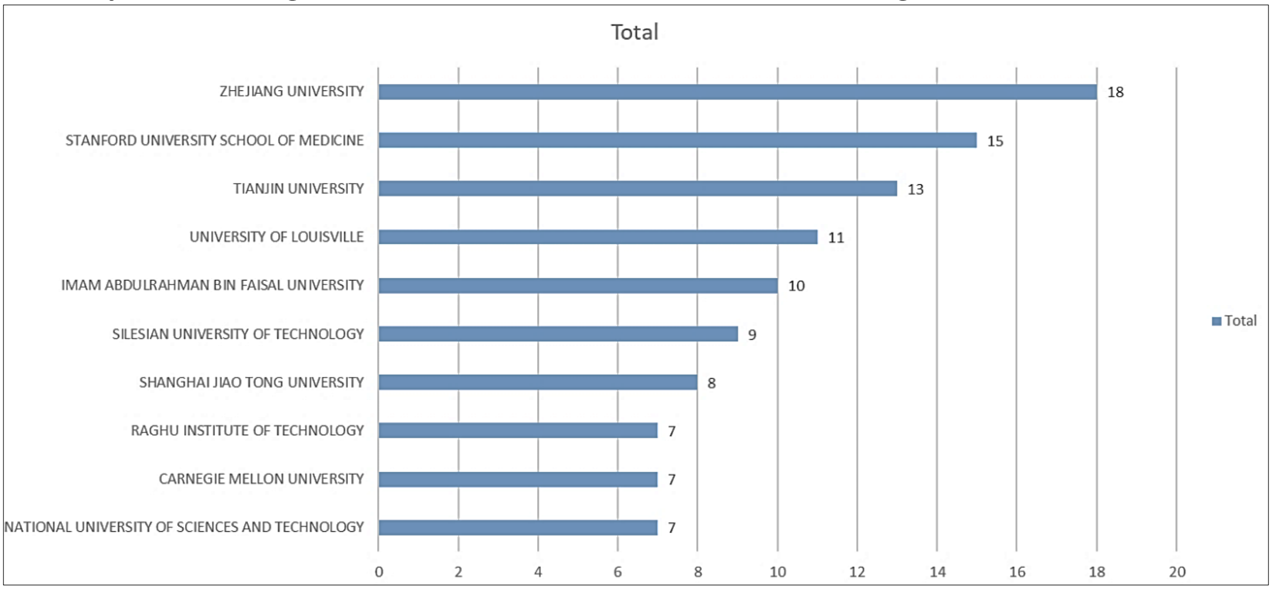


Figure 7. Top ten institutions based on the number of publications

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Insights of TDDBML

In this section, an in-depth examination of 41 research articles will be conducted, covering topics such as unbalanced data, thyroid disease, and machine learning. This review aims to provide insight into the concepts, methods, and potential future applications relevant to theorists and practitioners.

Thyroid disease kinds

As machine learning-based methods improved, scientists and doctors began using data-driven methods to determine if a patient had a thyroid issue from a blood sample. Patients, however, often have to wait until their symptoms have worsened before they see a doctor because of the difficulties involved in undertaking the numerous routine tests. However, ML-based methods enable early-stage diagnosis, which the subject himself can perform using inexpensive and compact sensors10 routinely. Thyroid disease can be classified into seven diagnostic categories: hypothyroidism, euthyroidism, goiter, thyroiditis, thyroid cancer, thyroid hormone resistance, and hyperthyroidism. There are two main types of hypothyroidism and hyperthyroidism. At least 15 of the papers out of the 41 chosen ones took into account two types because of their fatal consequences. Both types affect metabolism function, and severe conditions need medical attention[30].

Indeed, thyroid disease, particularly in its terminal stage, is related to an increased risk of cardiovascular illness, elevated blood pressure, higher cholesterol levels, and mental depression[31]. Therefore, to effectively treat patients with thyroid disease, it is crucial to diagnose the condition early. Ahmed et al. achieved a 98.2% accuracy rate in differentiating between hypothyroid and hyperthyroid states by training a deep neural network, as reported in their study[6]. Pal et al. compared the three machine learning models for predicting thyroid disease, including KNN, DT, and Multi-Layer Perceptron (MLP), and found that it achieves the highest accuracy of 94.23% [32]. The UCI thyroid disease open repository dataset was used in the study. On the other hand, Aljameel used an EANN-based approach to distinguish between thyroid cancer and non-cancer raw data using real-world data with 99% accuracy[33].

Figure 8 depicts the most commonly reported disease associated with the thyroid from cited studies. Thyroid disease has the largest cluster compared to other diseases related to machine learning techniques. In addition, several keywords have been repeated, which indicate that it is a technique used to predict a thyroid disease in the early stage, such as random forest, k-NN, and ANN.

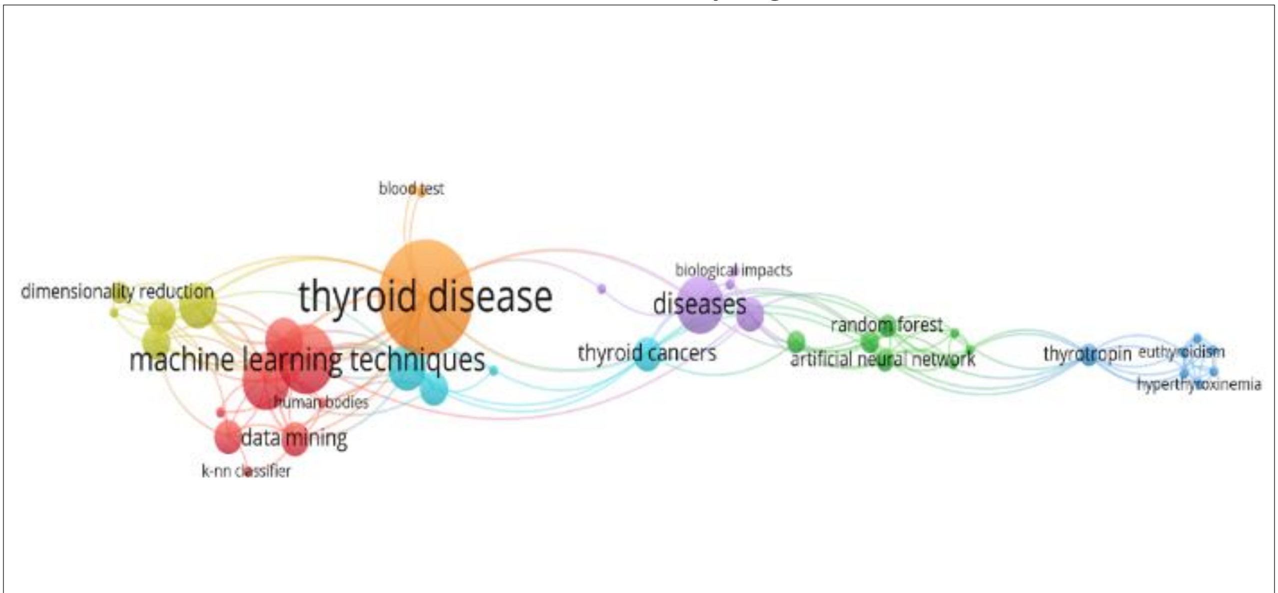


Figure 8. Illustration of most-reported thyroid disease (developed by Vosviewer software).

Machine learning algorithms

Table 4, illustrates that support vector machine (SVM) algorithms have received more

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

attention from researchers and practitioners than any other ML type in designing PTDBML models. At least 12 of the 41 studies that attempted to develop a model to diagnose thyroid disease used an SVM-based approach that used the standard technique in healthcare system prediction[34]. For instance, Płuciennik et al. has developed a model for thyroid cancer diagnostics, which achieved approximately 95% accuracy[35]. Vairale et al. compared SVM to Logistic Regression (LR), K-NN, and NN for identifying people with a hypothyroid disease on the actual case dataset. SVM showed the best performance among all algorithms, producing an accuracy of 99%[17].

On the Other side, the RF classifier is the following algorithm to enhance the thyroid disease prediction model: nine studies were conducted to develop a model for thyroid disease prediction. As a result, Alghamdi has designed an efficient predictive model to find thyroid cancer in the Prostate, Lung, Colorectal, and Ovarian (PLCO) dataset, defined as 155000 examples[36]. They used seven models the Logistic Regression model (LR), KNN, Ada boost classifier (AdaB), SVM, DT, Gaussian Naïve Bayes (GNB), RF, and Gradient Boosting classifier (GB); the RF has vital accuracy of 100%.

There are evident that, as time has progressed, a growing number of TDDBML model development efforts have focused on DL algorithms rather than classic ML. However, only 8 out of 41 studies focused on using DL to create a model for TDDBML, indicating that more research is needed. In order to classify individuals into normal, hyperthyroid, and hypothyroid categories, Guleria et al. used a thyroid cancer prediction system based on MLP. According to preliminary computational results, the proposed model identifies thyroid issues with an accuracy of 99.8% [37]. M Asif et al. proposed that MLP was the most effective algorithm, achieving an accuracy of 99.70% [38]. Zhou et al used ten ML algorithms through thyroid surgery to demonstrate a corresponding model. They employed a CNN model that utilized AUC and accuracy measures to identify patients at an early stage of thyroid disease. The study's main finding, based on data from 500 actual patients, was that the model achieved a 90% accuracy in accurately identifying individuals with thyroid disease, along with an AUC of 83%[39]. Other ML-based algorithms researchers use to create the TDDBML model include KNN[40], Hoeffding [41], XGBoost [31], and Adaboost and Bagging [42].

Imbalance challenges

One initial focus was tracking down previous research publications on thyroid disease that included analyses of imbalanced data. However, it became evident as one read through the articles that the vast majority of research either adopted data from other open sources or their studies used actual data and that in both situations, the datasets were unbalanced. As a result, the quality assessment revealed that eight articles relied on experimental results from the unbalanced dataset. In addition, recent studies have addressed the issue of imbalanced data's effect on model performances, which most studies ignored.

The imbalance problems are dealt with in various ways depending on the author. For instance, Zhou et al. assessed the model performance on unbalanced data classification by computing its f1 score, ROC-AUC curves, and accuracy rate [39]. N. Alghamdi has worked on the PLCO dataset, which shows patients that more classes have not been diagnosed with thyroid cancer, and fewer classes are diagnosed with thyroid cancer, and they relied on an under-sampling technique to handle imbalanced classes[36]. Aljameel S S et al, worked on a dataset with an imbalance (much more thyroid cancer cases than non-thyroid cancer cases); thus, they applied the SMOTEENN technique to avoid biasing the models toward one of the outcomes[33]. While [16,43,44] rely on SMOTE to handle imbalanced data issues to prevent bias in the performance measures. In cases where SMOTE is used to match the data, the overall model accuracy increases. Finally, Hayashi et al. have suggested a model like continuous Re-RX extract informative principles from the thyroid dataset with the correct values of subdivision rate for both the majority and minority classes [45].

Some researchers adopt DL-based solutions to replace all other algorithm-level methods. For instance, Selwal & Raoof used an MLP machine learning model to develop a more accurate system for predicting thyroid disease, which they tested on random samples of hyperthyroid, hypothyroid, and healthy subjects[37]. On the other hand, after choosing variables for thyroid illness prediction, several studies employ the Convolutional Neural Network (CNN), Long-Short-Term Memory (LSTM), and CNNLSTM. The authors demonstrate that their proposed model may achieve an AUC of 72%. Unfortunately, thyroid disease-related datasets are notoriously unbalanced, and few publications have investigated methods to address this issue outside of classification and fabricated models[18,23,46]. Many ways are available to handle imbalanced classes; however, few studies have been mentioned in this overview that impact model performance the most.

Results and Discussion

A thorough examination of 41 studies was done to understand the current practices and techniques for identifying thyroid disorders when working with an unbalanced dataset. The comprehensive analysis evaluated the following factors: thyroid disease type, applications, machine learning (ML) algorithms, and imbalance solutions. Based on the review of the literature, it was found that the authors employed a variety of datasets, including both real-case datasets and UCI datasets. Real-case datasets included data from medical institutions, such as blood test results and medical records. UCI datasets were obtained from the UC Irvine Machine Learning Repository, which includes various publicly available datasets. The authors also used dummy datasets created using various techniques, such as data augmentation and oversampling. The goal was to provide insights into the current practices and techniques used in thyroid disease diagnosis using an imbalanced dataset and to discuss the study's limitations and potential future research directions.

Overall, hypothyroidism and hyperthyroidism have received excellent attention in TDDBML. At the same time, other investigations looked into euthyroid[16] and thyroid surgery[39]. However, other types of thyroid diseases, such as thyroiditis, goiter, graves, and Hashimoto, a common problem among people suffering from malnutrition, trauma, surgery, or severe acute or chronic disease, received relatively less attention[47]. Most ML-based models are designed to detect thyroid disease patients and emphasize classification. Because of their availability and the issues connected with data imbalance concerns, most researchers heavily investigated popular datasets: UC Irvine Machine Learning Repository. However, a few research took into account real-world data[17,39,48] and large datasets[37,46]. A large amount of data helps the healthcare industry create more effective disease detection and decision-support systems[34]. The performance variation of the model has been detected in the study findings supplied as public source data and actual data. However, it cannot be denied that the performance of the models will be more accurate when the experiment is conducted using actual data [49]. In order to evaluate the effectiveness of ML-based models, it is required to use actual data rather than public repository data[50].

The instability of the model is one of the primary factors contributing to the restricted capabilities of the Clinical Decision Support Systems (CDSS) system[51]. Since clinical systems cannot function correctly using only old patient data, the CDSS model must be continuously refined and updated, considering new information [52]. Situations where it is necessary to collect data in real-time and train an ML model, such as the operating room during an emergency or a blood test conducted with the new devices, are likely to provide significant difficulties[53].

It is important to note that the majority of studies in the literature review utilized traditional machine learning methods such as SVM, RF, DT, KNN, and NB,[46,35,36] while fewer studies explored the use of deep learning algorithms such as ANN, MLP, CNN, LSTM, and BPNN[39,16,54,55]. While traditional machine learning algorithms have shown promising results in thyroid disease diagnosis, the potential of deep learning algorithms should not be overlooked, as they have shown success in various other medical applications. Among the traditional machine learning algorithms, the RF-based model has received the most attention from researchers[18, ,36, 48, 56, 57]. Due to its ability to handle thousands of variables and provide highly accurate classification[58]. However, it is essential to consider that different machine learning algorithms may perform differently depending on the specific characteristics of the dataset and the type of thyroid disease being diagnosed[59]. Therefore, it is crucial to explore a range of machine-learning algorithms and choose the most appropriate one for the problem.

The limited focus on handling imbalanced data in previous studies is a significant gap in the literature. Most studies have concentrated on feature selection techniques, neglecting the importance of handling imbalanced data, which can significantly impact the model's performance. This gap can be attributed to the lack of awareness among researchers about the impact of imbalanced data on the model's performance and the available techniques to handle it effectively. As a result, most of the models developed for thyroid disease diagnosis may not be suitable for real-world applications since they have not been tested with imbalanced data. Future researchers should focus on techniques for handling imbalanced data to address this gap. These techniques should be integral to model development and evaluation, and researchers should consider their impact on model performance. Techniques such as SMOTE[16,43,44], SMOTEENN[33], Under-sampling[36], and BRACID[45] have been used in previous studies to handle imbalanced data with varying degrees of success. Future researchers should consider these techniques and explore other advanced methods to improve the performance of the models.

One effective technique for handling imbalanced data is Cost-Sensitive Learning[60]. Cost-Sensitive Learning is a technique that assigns different misclassification costs to different classes. This technique can help to balance the misclassification costs between the minority and majority classes, thus improving the model's performance. Data Augmentation is another technique that can be used to address the imbalance problem[61]. Data Augmentation involves generating synthetic samples from the minority class to balance the class distribution. This technique can help to improve the performance of the model by increasing the diversity of the dataset. Threshold Adjustments are also effective in handling imbalanced data[62]. Threshold Adjustments involve adjusting the decision threshold of the model to improve the performance of the minority class.

Machine learning models often perform better intra-patient than inter-patient (inter-patient). Different data or patient characteristics can cause this[63]. For example, if a machine learning model is trained and evaluated on a dataset of individuals with one type of thyroid disease, it may not perform as well on a different dataset of individuals with a different type of thyroid disease. This could be due to patient-specific data, such as symptoms or blood test findings [64]. A larger diverse dataset of patients may be needed to train the machine learning model to generalize to a broader range of patient populations to increase its performance on inter-patient data. Choosing a different machine-learning method may be essential[65].

Most traditional classification methods try to find an ideal classifier that maximizes classification accuracy while keeping the misclassification cost constant, which can be problematic when dealing with imbalanced classes[66], considering the potential that the cost of misclassification may vary based on the probability distribution of the sample. In addition, most of the reported research included computationally costly techniques, including noising, thyroid segmentation, feature extractions, and classifications [67]. Deploying such a model in a real-world scenario could be challenging and an exciting avenue for future research. This aspect becomes even more critical when misclassification can lead to severe outcomes, as in the case of medical diagnostics.

In addition, there are challenges related to machine learning models' accuracy and reliability. Machine learning models are only as good as the data they are trained on, and if the data is of poor quality or biased, the model's predictions may not be accurate[68]. Therefore, it is essential to carefully evaluate the performance of machine learning models using appropriate evaluation metrics and consider the models' limitations when making predictions. Clinical diagnosis systems based on machine learning raise security problems for making diagnoses[69].

This study's findings significantly impact the development and evaluation of machine-learning models to diagnose and treat thyroid illness. The study emphasizes the need for additional research and development of machine learning models that can effectively handle imbalanced datasets, the importance of using real-world datasets, the potential of deep learning algorithms, and the challenges associated with deploying machine learning-based CDSS models. By addressing these consequences, researchers and doctors can use machine learning techniques to improve thyroid illness diagnosis and therapy accuracy and effectiveness. (

Table 4) collects the results of the studies cited in the literature to provide light on the ML-based prediction of thyroid disease.

Table 4. Literature Emphasized On TDDBML

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

References	Title	Algorithms	Imbalance	Evaluation	Dataset
35	Data Integration–Possibilities of Molecular and Clinical Data Fusion on the Example of Thyroid Cancer Diagnostics	SVM classifier (linear kernel)	NA	p-value below 0.05	200 real case
39	Predicting difficult airway intubation in thyroid surgery using multiple machine learning and deep learning algorithms	LR, RF, GB, XGB,LGBM,MLPC,GNB,CNN, LSTM, CNNLSTM	ROCNA AUC curves	92,91,91,91,91,90, 89,90, 89, 90	500 real case
48	Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset	RF, SVM, ANN, and NB	NA	90, 84, 88, 81	218 real case
36	Evaluation of classification models for predicting mortality rate using thyroid cancer data	LR, KN, SVC, GNB,DT, AdaB ,RF and GB	Under-sampling	98, 98, 99, 99,99, 99, 1, 1	155000
33	A Proactive Explainable Artificial Neural Network Model for the Early Diagnosis of Thyroid Cancer	explainable artificial neural network (EANN)	SMOTEENN	98	724 real case
16	Application of machine learning algorithms to Predict the thyroid disease risk: an experimental comparative study	ANN, CatBoost, XGB, RF, LGBM, DT, SVC, KNN ,GNB	SMOTE	95, 95, 95, 94,94, 94, 91, 89, 86	3,162 UCI
31	Thyroid Disease Prediction Using XGBoost Algorithms	LR, DT, KNN, and XGB	NA	81, 87, 96, 98	215 UCI
56	of thyroid disorder Effective voting ensemble of homogenous assembling with multiple attribute selection approaches for improved identification	DT, GB, LR, RF	NA	1	309 UCI
45	Use of the recursive rule extraction algorithm with continuous attributes to improve diagnostic accuracy in thyroid disease	Re RX	BRACID	96.70	7200 UCI
54	Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification	ELM	NA	99.68	12944 UCI
46	Developing an Explainable Machine Learning Based Thyroid Disease Prediction Model	LR	NA	91	215 UCI
40	Increasing the Prediction Accuracy for Thyroid Disease: A Step Towards Better Health for Society	KNN NN	NA	94, 98	3152 UCI
37	A Multi-layer perceptron-based intelligent thyroid disease prediction system.	MLP	NA	99	120 UCI
19	Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning	DT, ANN	NA	99, 99	3772 UCI
18	Constituent depletion and divination of hypothyroid prevalence using machine learning classification	RF, DT, NB, KNN, LR	NA	99	3164 UCI

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

References	Title	Algorithms	Imbalance	Evaluation	Dataset
70	Constructing a system for analysis of machine learning techniques for early detection of thyroid	ZeroR J48 Naïve bayes OneR	NA	60, 68, 41, 64	1,000 UCI
41	Decision tree ensemble techniques to predict thyroid disease	J48, RT, Hoeffding	NA	99, 97, 92	499 UCI
67	Prediction of Thyroid disease(Hypothyroid) in the Early Stage Using Feature Selection and Classification Techniques	SVM, DT, RF, LR, NB	NA	99, 99, 99, 96	519 UCI
43	Thy-Sys: A Preliminary Thyroid Wellness Assessment Through Machine Learning Using Pathological Factors	SVM, KNN, DT, SVMNA KNN	SMOTE	99.5	1464 UCI
71	Predictive Analysis for Thyroid Diseases Diagnosis Using Machine Learning	KNN, NB, DT	NA	92, 95, 99	1464 UCI
38	Computer-aided diagnosis of thyroid disease using machine learning algorithms	KNN, SVM, AdB, XGB, GPC, GBC,MLPC	NA	93, 96, 97, 96, 95, 98, 99	3164 UCI
72	A Machine Learning Approach to Predict Thyroid Disease at Early Stages of Diagnosis	DT, NB	NA	95	3000 UCI
57	Prediction of thyroid disorders using advanced machine learning techniques	NB, SVM, RF	NA	74, 92, 78	7200 UCI
17	Classification of Hypothyroid Disorder Using Optimized SVM Method	KNN, SVM, LR, NN	NA	97, 99, 95, 94	574 real case
73	A Study on Label TSH, T3, T4U, TT4, FTI in Hyperthyroidism and Hypothyroidism using Machine Learning Techniques	RF, SVM, KNN	NA	98, 97, 95	7200 UCI
55	Feature selection algorithms to improve thyroid disease diagnosis	MLP BPNN SVM ELM	NA	94, 95, 97, 98	215 UCI
32	Enhanced Prediction of Thyroid Disease Using Machine Learning Method	KNN, DT, MLP	NA	91, 94, 96	3163 UCI
44	Accuracy Assessment of Machine Learning Algorithm(s) in Thyroid Dysfunction Diagnosis	J48, MLP, NB, RF, SVM	SMOTE	99, 98, 98, 99, 98	4975 UCI
42	Efficient Thyroid Disease Prediction using Features Selection and Meta-Classifiers	Ada Boosting Bagging	NA	93, 99	774 UCI

Conclusion

This study uses imbalanced data to discover the most recent ML-based and data-driven developments and strategies in diagnosing thyroid disease. When developing ML-based systems for predicting thyroid disease in the real world, including real-patient data and using interpretable machine-learning methods to explain the final predictions is essential accurately. A comprehensive review of 41 papers suggests that more research is needed to prove reliable performance in healthcare settings. Although Deep Learning has come to dominate the area, SMOTE is still widely used as an Over-Sampling technique for handling unbalanced data by many academics and practitioners. Many researchers have noticed the development of an RF-based model for predicting thyroid disease since it is easier to train and can handle many features. Another big attraction is that they resist overfitting, making them useful in various machine-learning applications. The limits of ML that are discussed in the discussion sections may guide the direction of future research. Regardless, ML-based thyroid disease detection utilizing imbalanced data and innovative techniques is expected to uncover numerous undiscovered opportunities in the future.

Authors' Declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images, that are not ours, have been included with the necessary permission for re-publication, which is attached to the manuscript.
- Authors sign on ethical consideration's approval.
- Ethical Clearance: The project was approved by the local ethical committee in Universiti Teknologi Malaysia, Johor, Malaysia.

Authors' Contribution Statement

Dh S S: He conceived and designed the paper, which was his idea. He made the acquisition, analysis, and interpretation of the data. M Sh O: He did the Conception, design of the work and critical revision of the article. He did approve of the version being published.

References

1. Aschebrook-Kilfoy B, Ward MH, Sabra MM, Devesa SS. Thyroid Cancer Incidence Patterns in the United States by Histologic Type, 1992–2006. Thyroid. 2011 Feb; 21(2): 125–134. https://dx.doi.org/10.1089/thy.2010.0021
2. Harikrishna A, Ishak A, Ellinides A, Saad R, Christodoulou H, Spartalis E, et al. The impact of obesity and insulin resistance on thyroid cancer: A systematic review. Maturitas. 2019 Jul; 125: 45–49. https://dx.doi.org/10.1016/j.maturitas.2019.03.022
3. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. J Am Med Assoc. 2017; 317(13): 1338–1348. https://dx.doi.org/10.1001/jama.2017.2719
4. Clinic M. Hashimoto's disease. 2022; https://www.mayoclinic.org/diseases-conditions/hashimotos-disease/symptoms-causes/syc-20351855?_ga=2.22575032.287617514.1676515456-1166957844.1676515456
5. Thabit MA, Abdullah GH, AL-Rawi KF. Polymorphism study of MTHFR 677C?T and its correlation with oxidative stress and their influence on female infertility in Erbil – Iraq. Baghdad Sci J. 2017 Sep 3; 14(3): 611–618. https://dx.doi.org/10.21123/bsj.2017.14.3.0611
6. Ahmed I, Mohiuddin R, Muqeet MA, Kumar JA, Thaniserikaran A. Thyroid Cancer Detection using Deep Neural Network. Proc Int Conf Appl Artif Intell Comput 2022: 166–169. https://dx.doi.org/10.1109/ICAAIC53929.2022.9792854
7. Bahn RS, Burch HB, Cooper DS, Garber JR, Greenlee MC, Klein I, et al. Hyperthyroidism and Other Causes of Thyrotoxicosis: Management Guidelines of the American Thyroid Association and American Association of Clinical Endocrinologists. Thyroid. 2011 May; 17(3): 456–520. https://dx.doi.org/10.4158/EP.17.3.456
8. Porter KM, Ward M, Hughes CF, O'Kane M, Hoey L, McCann A, et al. Hyperglycemia and Metformin Use Are Associated With B Vitamin Deficiency and Cognitive Dysfunction in Older Adults. J Clin Endocrinol Metab. 2019 Oct 1; 104(10): 4837–4847. https://dx.doi.org/10.1210/jc.2018-01791
9. Papaleontiou M, Norton EC, Reyes-Gastelum D, Banerjee M, Haymart MR. Competing Causes of Death in Older Adults with Thyroid Cancer. Thyroid. 2021 Sep 1; 31(9): 1359–1365. https://dx.doi.org/10.1089/thy.2020.0929
10. Hamamurad QH, Jusoh NM, Ujang U. Modern City Issues, Management and the Critical Role of Information and Communication Technology. Int J Adv Comput Sci Appl. 2022; 13(4): 368–373. https://dx.doi.org/10.14569/IJACSA.2022.0130443
11. Leung AKC, Leung AAC. Evaluation and management of the child with hypothyroidism. World J Pediatr. 2019 Apr 8; 15(2): 124–134. https://dx.doi.org/10.1007/s12519-019-00230-w
12. Alexander EK, Pearce EN, Brent GA, Brown RS, Chen H, Dosiou C, et al. 2017 Guidelines of the American Thyroid Association for the Diagnosis and

Management of Thyroid Disease During Pregnancy and the Postpartum. Thyroid. 2017 Mar; 27(3): 315–389. https://dx.doi.org/10.1089/thy.2016.0457
13. Obschonka M, Audretsch DB. Artificial intelligence and big data in entrepreneurship: a new era has begun. Small Bus Econ. 2020; 55(3): 529–539. https://dx.doi.org/10.1007/s11187-019-00202-4
14. Garbuio M, Lin N. Artificial Intelligence as a Growth Engine for Health Care Startups: Emerging Business Models. Calif Manage Rev. 2019; 61(2): 59–83. https://dx.doi.org/10.1177/0008125618811931
15. Alfifi M, Shady M, Bataineh S, Mezher M. Enhanced Artificial Intelligence System for Diagnosing and Predicting Breast Cancer using Deep Learning. Int J Adv Comput Sci Appl. 2020; 11(7): 498–513. https://dx.doi.org/10.14569/IJACSA.2020.0110763
16. Islam SS, Haque MS, Miah MSU, Sarwar T Bin, Nugraha R. Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study. PeerJ Comput Sci. 2022 Mar 3; 8: e898. https://dx.doi.org/10.7717/peerj-cs.898
17. Vairale VS, Shukla S. Classification of Hypothyroid Disorder using Optimized SVM Method. Proc 2nd Int Conf Smart Syst Inven Technol. 2019: 258–263. https://dx.doi.org/10.1109/ICSSIT46314.2019.8987767
18. Devi MS, Shil A, Katyayan P, Surana T. Constituent Depletion and Divination of Hypothyroid Prevalance using Machine Learning Classification. Int J Innov Technol Explor Eng. 2019 Oct 30; 8(12): 1607–1612. https://dx.doi.org/10.35940/ijitee.L3150.1081219
19. Guleria K, Sharma S, Kumar S, Tiwari S. Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning. Meas Sensors. 2022 Dec; 24(5): 100482. https://dx.doi.org/10.1016/j.measen.2022.100482
20. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf. 2019; 28(3): 231–237. https://dx.doi.org/10.1136/bmjqs-2018-008370
21. Ahsan MM, Ahad MT, Soma FA, Paul S, Chowdhury A, Luna SA, et al. Detecting SARS-CoV-2 from chest X-Ray using artificial intelligence. Ieee Access. 2021; 9: 35501–35513. https://dx.doi.org/10.1109/ACCESS.2021.3061621
22. Feng W, Huang W, Ren J. Class Imbalance Ensemble Learning Based on the Margin Theory. Appl Sci. 2018 May 18; 8(5):815. https://dx.doi.org/10.3390/app8050815
23. Lee KS, Park H. Machine learning on thyroid disease: a review. Front Biosci - Landmark. 2022; 27(3): 101. https://dx.doi.org/10.31083/j.fbl2703101
24. Mendoza AM, Hernandez RM. Application of Data Mining Techniques in Diagnosing Various Thyroid Ailments: A Review. In: 13 th Int Conf Inf Commun Technol Syst. 2021: 207–212. https://dx.doi.org/10.1109/ICTS52701.2021.9608400
25. Anari S, Tataei Sarshar N, Mahjoori N, Dorosti S, Rezaie A. Review of Deep Learning Approaches for Thyroid Cancer Diagnosis. Math Probl Eng. 2022: 1–8. https://dx.doi.org/10.1155/2022/5052435
26. Okoli C, Schabram K. A Guide to Conducting a Systematic Literature Review of Information Systems Research. Soc Sci Res Netw. 2010; 10(26): 1–51. https://dx.doi.org/10.2139/ssrn.1954824
27. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021; 372(2021): n71. https://dx.doi.org/10.1136/bmj.n71
28. Fahimnia B, Sarkis J, Davarzani H. Green supply chain management: A review and bibliometric analysis. Int J Prod Econ. 2015 Apr; 162(2015): 101–114. https://dx.doi.org/10.1016/j.ijpe.2015.01.003
29. Gregg W, Erdmann C, Paglione L, Schneider J, Dean C. A literature review of scholarly communications metadata. Res Ideas Outcomes. 2019 Aug; 5: e38698. https://dx.doi.org/10.3897/rio.5.e38698
30. Abeer Uthman Moosa FAAMAM. Effect of Hyper and Hypothyroidism on Lipid Profile and Liver Function of Male Rats. Baghdad Sci J. 2011; 8(4): 926–933. https://dx.doi.org/10.21123/bsj.8.4.926-933
31. Sankar S, Potti A, Naga Chandrika G, Ramasubbareddy S. Thyroid Disease Prediction Using XGBoost Algorithms. J Mob Multimed. 2022; 18(3): 917–934. https://dx.doi.org/10.13052/jmm1550-4646.18322
32. Pal M, Parija S, Panda G. Enhanced Prediction of Thyroid Disease Using Machine Learning Method. Proc IEEE VLSI DCS 2022 3rd IEEE Conf VLSI Device, Circuit Syst. 2022(February): 199–204. https://dx.doi.org/10.1109/VLSIDCS53788.2022.9811472
33. Aljameel SS. A Proactive Explainable Artificial Neural Network Model for the Early Diagnosis of Thyroid Cancer. Comput. 2022; 10(10): 183. https://dx.doi.org/10.3390/computation10100183
34. Kamra V, Kumar P, Mohammadian M. Diagnosis support system for general diseases by implementing a novel machine learning based classifier. Int J Comput Digit Syst. 2021; 10(1): 737–746. https://dx.doi.org/10.12785/ijcds/100168
35. Płuciennik A, Płaczek A, Wilk A, Student S, Oczko-Wojciechowska M, Fujarewicz K. Data Integration–Possibilities of Molecular and Clinical Data Fusion on the Example of Thyroid Cancer Diagnostics. Int J Mol Sci. 2022; 23(19): 11880. https://dx.doi.org/10.3390/ijms231911880
36. Alghamdi NS. Evaluation of classification models for predicting mortality rate using thyroid cancer data. J Comput Sci. 2019; 15(1): 131–142. https://dx.doi.org/10.3844/jcssp.2019.131.142
37. Selwal A, Raoof I. A Multi-layer perceptron based intelligent thyroid disease prediction system. Indones

J Electr Eng Comput Sci. 2020; 17(1): 524–532. https://dx.doi.org/10.11591/ijeecs.v17.i1.pp524-532

38. Asif MAAR, Nishat MM, Faisal F, Shikder MF, Udoy MH, Dip RR, et al. Computer aided diagnosis of thyroid disease using machine learning algorithms. Proc 2020 11th Int Conf Electr Comput Eng. 2020; 4: 222–225. https://dx.doi.org/10.1109/ICECE51571.2020.9393054
39. Zhou CM, Wang Y, Xue Q, Yang JJ, Zhu Y. Predicting difficult airway intubation in thyroid surgery using multiple machine learning and deep learning algorithms. Front Public Heal. 2022; 10(2022): 1–14. https://dx.doi.org/10.3389/fpubh.2022.937471
40. Jha R, Bhattacharjee V, Mustafi A. Increasing the Prediction Accuracy for Thyroid Disease: A Step Towards Better Health for Society. Wirel Pers Commun. 2022; 122(2): 1921–1938. https://dx.doi.org/10.1007/s11277-021-08974-3
41. Yadav DC, Pal S. Decision tree ensemble techniques to predict thyroid disease. Int J Recent Technol Eng. 2019; 8(3): 8242–8246. https://dx.doi.org/10.35940/ijrte.C6727.098319
42. Priyadharsini D, Sasikala S. Efficient Thyroid Disease Prediction using Features Selection and Meta-Classifiers. Proc 6th Int Conf Comput Methodol Commun. 2022 :1236–1243. https://dx.doi.org/10.1109/ICCMC53470.2022.9753986
43. Francisco IR, Ferolin MBJ, Pena CF, Ferolin RJ. Thy-Sys: A Preliminary Thyroid Wellness Assessment Through Machine Learning Using Pathological Factors. Proc 1st Int Conf Inf Comput Res iCORE 2021: 44–49. https://dx.doi.org/10.1109/iCORE54267.2021.00027
44. Danjuma KJ, Maksha Wajiga G, Garba EJ, Sandra Ahmadu A, Longe OB. Accuracy Assessment of Machine Learning Algorithm(s) in Thyroid Dysfunction Diagnosis. Proc 2022 IEEE Niger 4th Int Conf Disruptive Technol Sustain Dev, NIGERCON. 2022: 1–5. https://dx.doi.org/10.1109/NIGERCON54645.2022.9803113
45. Hayashi Y, Nakano S, Fujisawa S. Use of the recursive-rule extraction algorithm with continuous attributes to improve diagnostic accuracy in thyroid disease. Inform Med Unlocked. 2015; 1(2015): 1–8. https://dx.doi.org/10.1016/j.imu.2015.12.003
46. Arjaria SK, Rathore AS, Chaubey G. Developing an Explainable Machine Learning-Based Thyroid Disease Prediction Model. Int J Bus Anal. 2022; 9(3): 1–18. https://dx.doi.org/10.4018/IJBAN.292058
47. Vanderpump MPJ. The epidemiology of thyroid disease. Br Med Bull. 2011; 99(1): 39–51. https://dx.doi.org/10.1093/bmb/ldr030
48. Olatunji SO, Alotaibi S, Almutairi E, Alrabae Z, Almajid Y, Altabee R, et al. Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset. Comput Biol Med. 2021; 131(April): 104267. https://dx.doi.org/10.1016/j.compbiomed.2021.104267
49. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. BMC Med Res Methodol. 2022 Nov 5; 22(1): 287. https://dx.doi.org/10.1186/s12874-022-01768-6
50. Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci. 2021 May 22; 2(3): 160. https://dx.doi.org/10.1007/s42979-021-00592-x
51. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. Br Med J. 2005; 330(7494):765–768. https://dx.doi.org/10.1136/bmj.38398.500764.8f
52. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One. 2017; 12(4): e0174944. https://dx.doi.org/10.1371/journal.pone.0174944
53. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med. 2019 Sep 19; 25(9): 1337–1340. https://dx.doi.org/10.1038/s41591-019-0548-6
54. Juneja K. Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification. Wirel Pers Commun. 2022 Sep 1; 126(2): 1805–1842. https://dx.doi.org/10.1007/s11277-022-09823-7
55. Pavya K, Srinivasan B. Feature selection algorithms to improve thyroid disease diagnosis. IEEE Int Conf Innov Green Energy Healthc Technol. 2017: 1–5. https://dx.doi.org/10.1109/IGEHT.2017.8094070
56. Akhtar T, Gilani SO, Mushtaq Z, Arif S, Jamil M, Ayaz Y, et al. Effective voting ensemble of homogenous ensembling with multiple attribute-selection approaches for improved identification of thyroid disorder. Electron. 2021; 10(23): 3026. https://dx.doi.org/10.3390/electronics10233026
57. Duggal P, Shukla S. Prediction of thyroid disorders using advanced machine learning techniques. Proc Conflu 2020 - 10th Int Conf Cloud Comput Data Sci Eng. 2020; 670–675. https://dx.doi.org/10.1109/Confluence47617.2020.9058102
58. Deng W, Huang Z, Zhang J, Xu J. A Data Mining Based System for Transaction Fraud Detection. 2021 IEEE Int Conf Consum Electron Comput Eng. 2021: 542–545. https://dx.doi.org/10.1109/ICCECE51280.2021.9342376
59. Fenza G, Gallo M, Loia V, Orciuoli F, Herrera-Viedma E. Data set quality in Machine Learning: Consistency measure based on Group Decision

Making. Appl Soft Comput. 2021 Jul; 106(9): 107366. https://dx.doi.org/10.1016/j.asoc.2021.107366
60. Sun Y, Wong AKC, Kamel MS. Classification of imbalanced data: A review. Int J Pattern Recognit Artif Intell. 2009; 23(4): 687–719. https://dx.doi.org/10.1142/S0218001409007326
61. Shamsolmoali P, Zareapoor M, Shen L, Sadka AH, Yang J. Imbalanced data learning by minority class augmentation using capsule adversarial networks. Neurocomputing. 2021; 459: 481–493. https://dx.doi.org/10.1016/j.neucom.2020.01.119
62. Esposito C, Landrum GA, Schneider N, Stiefl N, Riniker S. GHOST: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning. J Chem Inf Model. 2021; 61(6): 2623–2640. https://dx.doi.org/10.1021/acs.jcim.1c00160
63. Kann BH, Hosny A, Aerts HJWL. Artificial intelligence for clinical oncology. Cancer Cell. 2021; 39(7): 916–927. https://dx.doi.org/10.1016/j.ccell.2021.04.002
64. Paleyes A, Urma RG, Lawrence ND. Challenges in Deploying Machine Learning: A Survey of Case Studies. ACM Comput Surv. 2022; 55(6):1–29. https://dx.doi.org/10.1145/3533378
65. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019; 380(14): 1347–1358. https://dx.doi.org/10.1056/nejmra1814259
66. Gan D, Shen J, An B, Xu M, Liu N. Integrating TANBN with cost sensitive classification algorithm for imbalanced data in medical diagnosis. Comput Ind Eng. 2020 Feb; 140(2): 106266. https://dx.doi.org/10.1016/j.cie.2019.106266
67. Riajuliislam M, Rahim KZ, Mahmud A. Prediction of Thyroid Disease(Hypothyroid) in Early Stage Using Feature Selection and Classification Techniques. Int Conf Inf Commun Technol Sustain Dev. IEEE. 2021. p. 60–64. https://dx.doi.org/10.1109/ICICT4SD50815.2021.9397052
68. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. npj Digit Med. 2018; 1(1): 1–10. https://dx.doi.org/10.1038/s41746-018-0029-1
69. Liang H, Tsui BY, Ni H, Valentim CCS, Baxter SL, Liu G, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med. 2019 Mar 11; 25(3): 433–438. https://dx.doi.org/10.1038/s41591-018-0335-9
70. Rasheeduddin S, Rajasekhar Rao K. Constructing a system for analysis of machine learning techniques for early detection of thyroid. Int J Eng Adv Technol. 2019; 8: 1978–1981. https://dx.doi.org/10.35940/ijeat.F1385.0986S319
71. Peya ZJ, Chumki MKN, Zaman KM. Predictive Analysis for Thyroid Diseases Diagnosis Using Machine Learning. 2021 Int Conf Sci Contemp Technol. 2021: 4–9. https://dx.doi.org/10.1109/ICSCT53883.2021.9642544
72. Rao AR, Renuka BS. A Machine Learning Approach to Predict Thyroid Disease at Early Stages of Diagnosis. 2020 IEEE Int Conf Innov Technol. 2020; 1–4. https://dx.doi.org/10.1109/INOCON50539.2020.9298252
73. Shahid AH, Singh MP, Raj RK, Suman R, Jawaid D, Alam M. A Study on Label TSH, T3, T4U, TT4, FTI in Hyperthyroidism and Hypothyroidism using Machine Learning Techniques. 2019 Int Conf Commun Electron Syst. 2019;(Icces): 930–933. https://dx.doi.org/10.1109/ICCES45898.2019.9002284

Published Online First: August, 2023
https://dx.doi.org/10.21123/bsj.2023.8544
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

تشخيص أمراض الغدة الدرقية على أساس التعلم الآلي: مراجعة منهجية للأدبيات

ذكري صابر صالح، محمد شهيزان عثمان

كلية الحاسبات ، جامعة التكنولوجيا ماليزيا ، جوهور ، ماليزيا.

الخلاصة

مرض الغدة الدرقية مرض شائع يصيب الملايين في جميع أنحاء العالم. ومن الممكن أن يساعد التشخيص والعلاج المبكر لمرض الغدة الدرقية لمنع حدوث المضاعفات أكثر خطورة وتحسين النتائج الصحية على المدى الطويل. ومع ذلك ، يمكن أن يكون تشخيص مرض الغدة الدرقية صعبًا بسبب أعراضه المتغيرة والاختبارات التشخيصية المحدودة. ومن خلال معالجة كميات هائلة من البيانات ورؤية الاتجاهات التي قد لا تكون واضحة على الفور للأطباء البشريين ، قد تكون خوارزميات التعلم الآلي قادرة على زيادة الدقة في تشخيص مرض الغدة الدرقية. تُستخدم استراتيجية مراجعة منهجية للادب في هذه الدراسة لإعطاء نظرة عامة شاملة عن الأدبيات الموجودة حول التنبؤ بالبيانات المتعلقة بأمراض الغدة الدرقية التي تم تشخيصها باستخدام التعلم الالي .وتتضمن هذه الدراسة 168 مقالة منشورة بين عامي 2013 و 2022 ، تم جمعها من المجلات عالية الجودة والتحليل التلوي التطبيقي. كانت فئة تشخيص أمراض الغدة الدرقية ، والتقنيات ، والتطبيقات ، والحلول من بين العديد من العناصر التي تم النظر فيها والبحث عنها عند مراجعة 41 مقالة من الأدبيات المذكورة المستخدمة في هذا البحث. استنادًا إلى نتائج منهجية للأدب الخاص بنا ، فإن التطبيق الفعلي للتقنية الحالية وفعاليتها مقيدان بالعديد من المشكلات المعلقة المرتبطة بعدم التوازن. في تشخيص امراض الغدة الدرقية، تزيد تقنية التعليم الألي من اتخاذ القرار المستند إلى البيانات. في التحليل التلوي ،كما تمت معالجة 168 وثيقة ، و تضمين 41 وثيقة عن تشخيص امراض الغدة الدرقية لتحليل المراقبة. قد توجه حدود التعلم الالي التي تمت مناقشتها في أقسام المختصة اتجاه البحث المستقبلي. يتنبأ هذا التحليل بأن اكتشاف مرض الغدة الدرقية المستند إلى التعلم الآلي مع البيانات غير المتوازنة وغيرها من الأساليب الجديدة قد يكشف عن العديد من الاحتمالات غير المحققة في المستقبل. تسعى هذه الدراسة إلى اكتشاف أحدث التطورات والاستراتيجيات المستندة إلى التعلم الآلي والمستندة إلى البيانات في تشخيص أمراض الغدة الدرقية باستخدام غير متوازن.

الكلمات المفتاحية: أمراض الغدة الدرقية، التعلم الآلي ، التعلم العميق ، البيانات غير المتوازنة ، التصنيف.