**Thyroid disease diagnosis based on machine learning: A systematic literature review**

**Dhekre Saber Saleh1, Shahizan Othman Mohd1, Qasim Hamakhurshid Hamamurad2, Wshyar Omar Khudhur3,**

*1 School of Computing, Universiti Teknologi Malaysia, Johor, Malaysia*

*2* *Information System*, *Universiti Teknologi Malaysia, Johor, Malaysia*

*3* *Koya Technical Institute – Erbil Polytechnic University (EPU), Erbil, Iraq*

*…*

*NDepartment,Name of the Organization, City, Country*

*E-mail address: author1@email.com, author2@email.com, …, authorN@email.com*

*Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##*

**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. 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 imbalanced and high-dimensional data. 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, we predict that ML-based thyroid disease detection with imbalanced data, dimensionality reduction, and other novel approaches may reveal numerous unrealized possibilities in the near future. This study seeks to discover the most recent ML-based and data-driven developments and strategies in diagnosing thyroid disease using imbalanced and high-dimensional data.

Keywords: Machine learning, Thyroid disease, Deep learning, High dimensional, Imbalanced data

# **Introduction**

Thyroid disease affects the thyroid gland, which produces hormones that balance the metabolism. Thyroid disease can be classified into two primary categories: hypothyroidism and hyperthyroidism. Hypothyroidism is defined by a thyroid hormone reduction and can lead to signs such as weight gain, tiredness, and constipation. Hyperthyroidism is distinguished by an overload of thyroid hormones and can lead to signs such as irritability, weight loss, and tremors. The Mayo Clinic states that thyroid disease has various causes, including genetics, autoimmune conditions, and exposure to radiation[1]. It can also be caused by certain medications or problems with the pituitary gland, which regulates the thyroid. Thyroid diseases affect an estimated 400 million population annually, making them a significant source of disability, according to the information provided by the World Health Organization (WHO). Despite this, the WHO does not track the number of people who die from thyroid cancer[2]. However, early diagnosis of the thyroid and immediate treatment can decrease deaths[3]. There are several ways to identify thyroid diseases, such as physical examination, and blood tests, enabling a more accurate diagnosis[4].

Ultrasound, thyroid function tests, and thyroid biopsy. Diagnosing thyroid disease can be difficult for healthcare providers because of some reasons, such as having symptoms similar to other illnesses, experiencing varying symptoms, having limited access to specialized care, and having limited diagnostic tests available[5], [6].

With the advancement of machine learning (ML) in the healthcare industry, many experts in the field have deemed Thyroid Disease Diagnosis Based on Machine Learning (TDDBML) as a viable option. ML improves disease diagnosis accuracy and efficiency[7]; algorithms can scan high amounts of data and recognize patterns that doctors may miss, enabling a more accurate diagnosis [7].

ML algorithms can evaluate electronic health data and patient monitoring devices to identify early disease indications and avert consequences, improving patient care and reducing wait times[8]. Consequently, numerous investigations utilizing various thyroid illness datasets proposed TDDBML. Be a case in point, S Islam et al. used various ML approaches, including artificial neural network (ANN), CatBoost, XGBoost, random forest (RF), LigthGBM, decision tree (DT), support vector classifier (SVC), K-Nearest Neighbors (KNN), and GaussianNB. Based on their data, the ANN classifier was able to construct a model that focuses on the prediction analysis of thyroid diseases with a 96% accuracy rate[9]. Later, Pluciennik et al. introduced a prediction model for a thyroid disease that combines molecular and clinical data using ML techniques with SVC (linear kernel) and enabling a more accurate diagnosis Shyamala Devi et al. developed a model for predicting hypothyroid disease using various machine learning techniques such as Naive Bayes (NB), Linear Regression (LR), Random Forest (RF), Decision Tree (DT), and KNN, and it achieved an accuracy rate of 99%[11]. Lastly, Guleria et al. created a model that used ANN (Artificial Neural Network) to predict hypothyroidism early on and achieved 100% accuracy[12].

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 makes it hard for doctors to evaluate the algorithm's diagnosis and raises bias and reliability concerns[13]. For instance, DLs have numerous invisible layers, but it is not always easy to tell what role each plays in the model's predictions[14]. 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 total value[15]. 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 I summarizes review studies investigating the use of ML techniques in TDD; it shows that the SLR is mainly ignored in support of the ML techniques. For example, K Lee et al. presented an SLR whose main issue was that machine learning methods vary for data utilized in diagnosing thyroid disease[16]. However, most existing thyroid disease datasets have imbalanced data, so examining how well ML performs on such data is also necessary. The research did not offer time periods[17]. With the growth of ML-based diagnosis, SLR with meta-analysis is expected to fill the gap between existing studies.

The growing number of papers in TDDBML highlights the need to draw upon the existing body of knowledge to generate novel insights and establish new lines of inquiry. SLR from Scopus and WoS databases. A total of 168 papers were used for the metadata analysis, with further examination being performed on 41 of those.

The primary objectives of metadata analysis are to identify leading academic institutions, essential research subjects, and high-quality sources. What are the most frequently used words in thyroid disease diagnoses depending on ML? Who are the prospective authors in the field? How much study is integrated into the findings? The detailed examination of 41 publications answered the following research objective: What are the existing ML and DL-based approaches for diagnosing thyroid disease? What are the recent techniques for dealing with datasets with an imbalanced class ratio? What are the current approaches to handling high dimensional?

TABLE I RELATED RESEARCH FOR THYROID DISEASE PREDICTION BASED ML

| **Paper title** | **Date Range** | **Study focus** | **Algorithm** | **Imbalance challenge** | **Feature challenge** | **Evaluation metrics** | **Meta-analysis** | **Content analysis** | **SLR** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Machine learning on thyroid disease: a review [16] | 2020 or later | ML algorithms differ for data. | ✓ |  |  |  |  |  |  |
| Application of Data Mining Techniques in Diagnosing Various Thyroid Ailments: A Review [17] | Not specified | data mining methods | ✓ |  | ✓ |  |  |  |  |
| review of Deep Learning Approaches for Thyroid Cancer Diagnosis [18] | 2018 or late | Estimating diagnostic accuracy of deep learning | ✓ |  |  |  |  |  |  |
| Our study | 2013 to 2022 | thyroid disease diagnosis based on machine learning | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

The SLR's true purpose is to operate as a resource for both researchers and practitioners by not only summarizing the most up-to-date methods and developments in the field but also by pinpointing the gaps in knowledge that may be bridged by the creation of a more sophisticated TDDBML model the remaining parts of the article are organized as described below: The SLR method is briefly labeled in Section 2, and Section 3 presents the results and discussion. Finally, Section 4 recaps our findings and Section 5 conclusions.

# **Method**

An SLR positions research questions before systematically searching for, selecting, and evaluating studies to see what information may be obtained from them[19]. This method is selected because it offers an accurate and dependable way to synthesize academic material and is generally acknowledged in various studies disciplines. The eligible studies items for meta-analyses, and systematic reviews or PRISMA, guidelines are followed while presenting the SLR.

## 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 [20]. The search covers 2013 to 2022 and includes all essential papers published during this time. We utilized terms such as "thyroid", "machine learning", "imbalance", "high dimensional", and "deep learning" to find relevant publications. Boolean operators and various keywords are used to improve the search.

## Screening initial data and determining eligibility

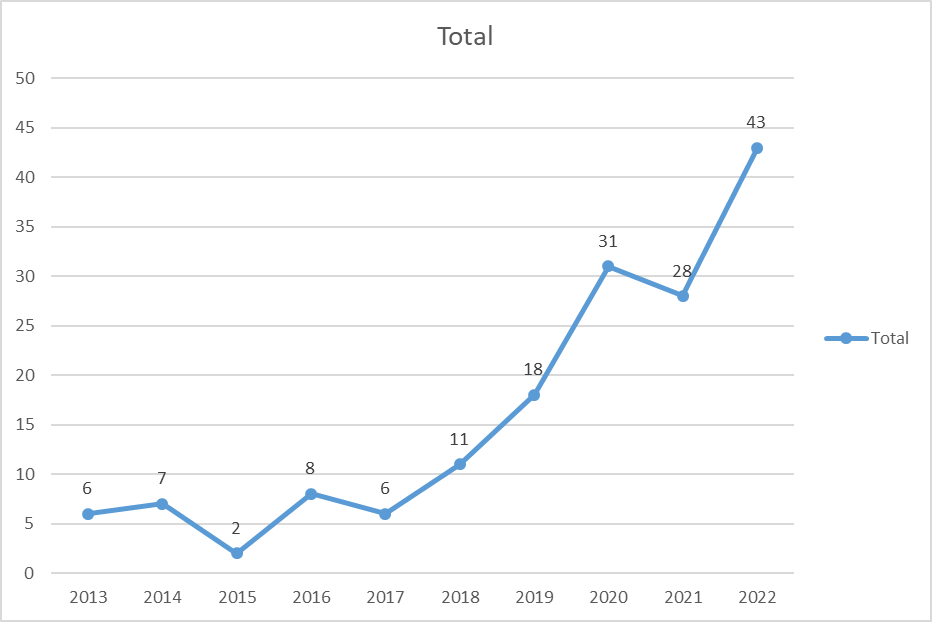
The process of searching for articles related to using ML techniques in diagnosing thyroid disease. The initial search using specific keywords on the Scopus database returned 2,182 articles and 486 articles on the WoS database. However, after applying the year limits of 2013-November-2022, the number of papers was narrowed down to 1159. Then 168 articles were selected for further screening based on document type, language, subject area, and keyword restrictions. The information for these 168 articles was exported as a CSV file in Excel, eliminating duplicates or leaving blanks.

Besides, we reviewed the titles and abstracts of the remaining 168 unique papers. A standardized extraction form was used to extract the most relevant information. Research unrelated to machine learning but focused on thyroid disease was excluded. The researchers also decided not to include book chapters, ultrasound imaging, non-human studies, or reviews in their analysis. Fig. 1 shows that all inclusion criteria were met by the 41 full-text publications that made it through the initial title and abstract screening.

# **Observations and findings**

The following section will discuss the findings and insights gained by analyzing the metadata.

These results are based on a meta-study of 168 papers, an analysis of their corresponding metadata, and a content analysis of 41 publications.



## Metadata analysis

Metadata analysis helps understand scholarly literature by extracting information about the scholarly process's authors, articles, journals, and other elements [21]. We used it for 168 papers in the metadata analysis and categorized them according to year, publication, publisher, country, topic, financing, and institution.

Fig. 1 PRISMA approach utilized in this research



### Published by year

As shown in Fig. 2, 168 papers were reviewed to see how many dealt with thyroid disease prediction using ML algorithms over the past decade. Publishing was expanding at a constant rate, and it is expected that this increase will be intensified significantly in the years 2020 and 2022. For example, in 2022, there were around 43 new papers published; in 2020, there were just 31 new papers 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 side, we can note the tiny number of papers published, particularly from 2013 to 2017, when there were just a few papers. Consequently, increasing focus and concern are directed toward diagnosing thyroid disease, including classification issues and other data-driven concerns.

Fig. 2 published papers on thyroid disease by using ML.

### Most Relevant Authors

According to Fig 3, Fu C and Liu W. have written the most relevant papers whole five and are thus the most influential writers. We did a similar data analysis to track the writers output over time. We discovered that Fu C and Liu W combined four 2021 articles that were cited 10.5 times. Fig. 4 displays the annual citation total of 62.57 for Chen H's essential citation 2016, which consists of 1 item published in 2016.

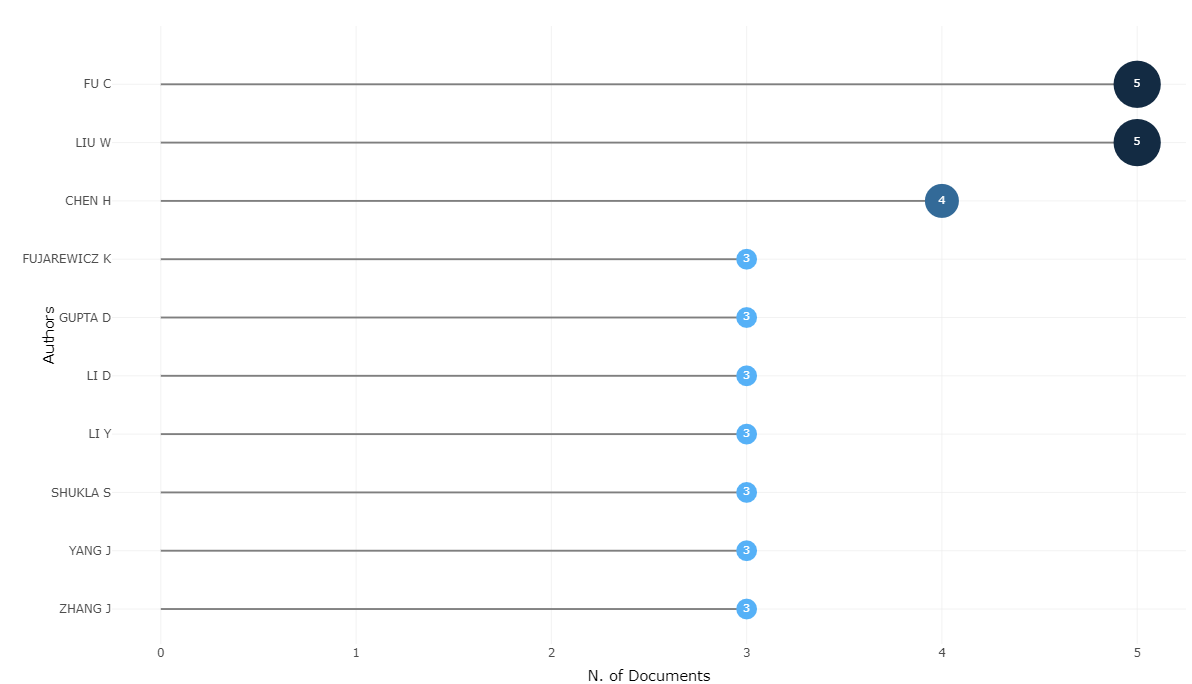
****

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

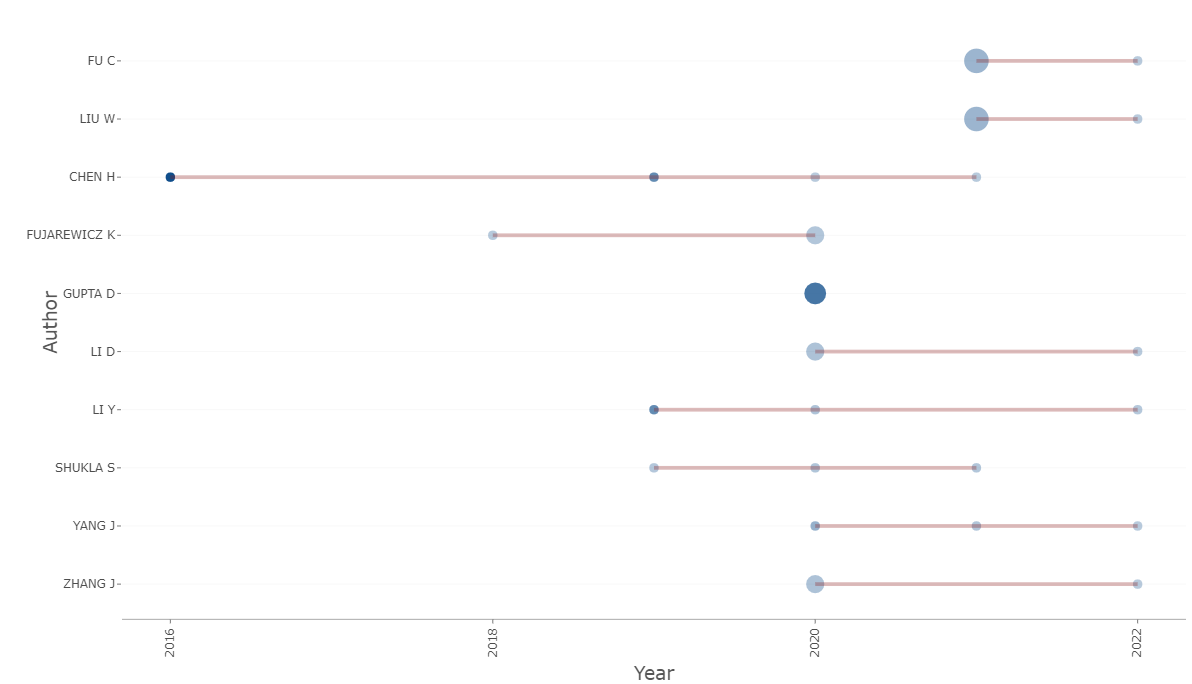
****

Fig. 4 Authors' production over time.

### Most Relevant Sources

As shown in Fig. 5, 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.

TABLE II MOST FREQUENTLY UTILIZED 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 |
| 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 |

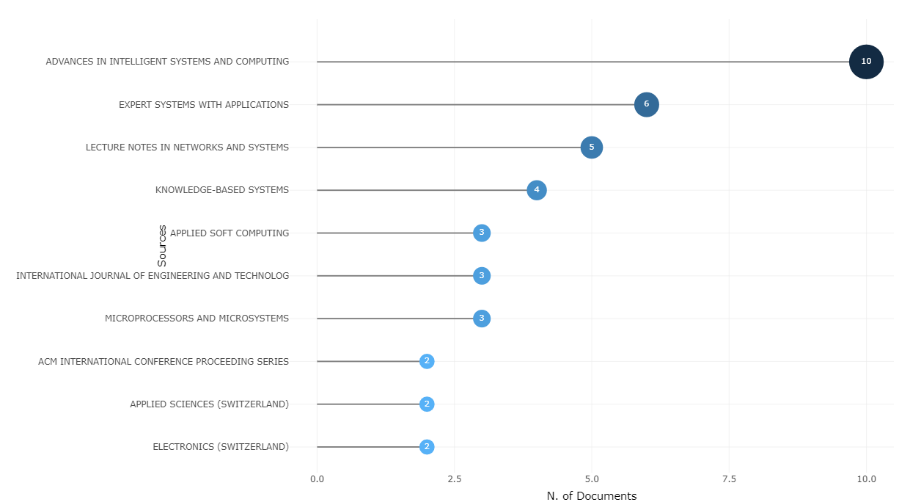


Fig. 5 Most relevant sources. The x-axis shows the number of documents; the y-axis indicates

### Most frequently words used in the titles and keywords Table

The most common single, double, and triple keywords that were employed in the paper titles are shown in table II. The R-software program is used to identify the most popular keywords. Even though our main goal was to find and analyze articles that only focused on concepts like machine learning, deep learning, features, imbalance class, and thyroid disease, we were surprised to find that only "thyroid" and "disease" ranked among the top most

TABLE III 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 |

When we looked specifically at the keywords used by the authors in the articles' keyword sections, however, we came across some interesting results. The most popular terms used in papers' keywords are listed in table II. 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. Articles often use the exact phrases which are listed below.

To find the most general terms in a complicated setting, a word cloud is a simple way to detect the common themes and keywords that are utilized in the referenced articles. Software-generated word clouds are shown in Fig. 6, with larger and bolder fonts showing the most often-used words and smaller and more common fonts highlighting the less frequently used phrases.

### Trending Topics

Our research allowed us to identify the prevailing themes concealed within the dataset's titles and abstracts. There were 68 occurrences of the word "thyroid" in the titles. Fig. 7 displays that "classification" was the second most popular theme in the literature, appearing 42 times. Fig. 7 displays the most famous abstract themes, which were "thyroid" (492 occurrences) and "data" (234 occurrences). Classification, sickness, algorithms, and so on were other common themes in the abstracts.

### Publication by institutions

Fig. 8 shows the most cited articles sorted by the institution where their authors work. Regarding the number of publications related to thyroid disease, the Figure indicates that the leader is the Hefei University of Technology in China. In total, 14 articles came out of the institutes. Then the second highest number of publications came from China's Jilin University, with 13.

Fig. 6 Word cloud for most frequently used keywords in thyroid disease publications.

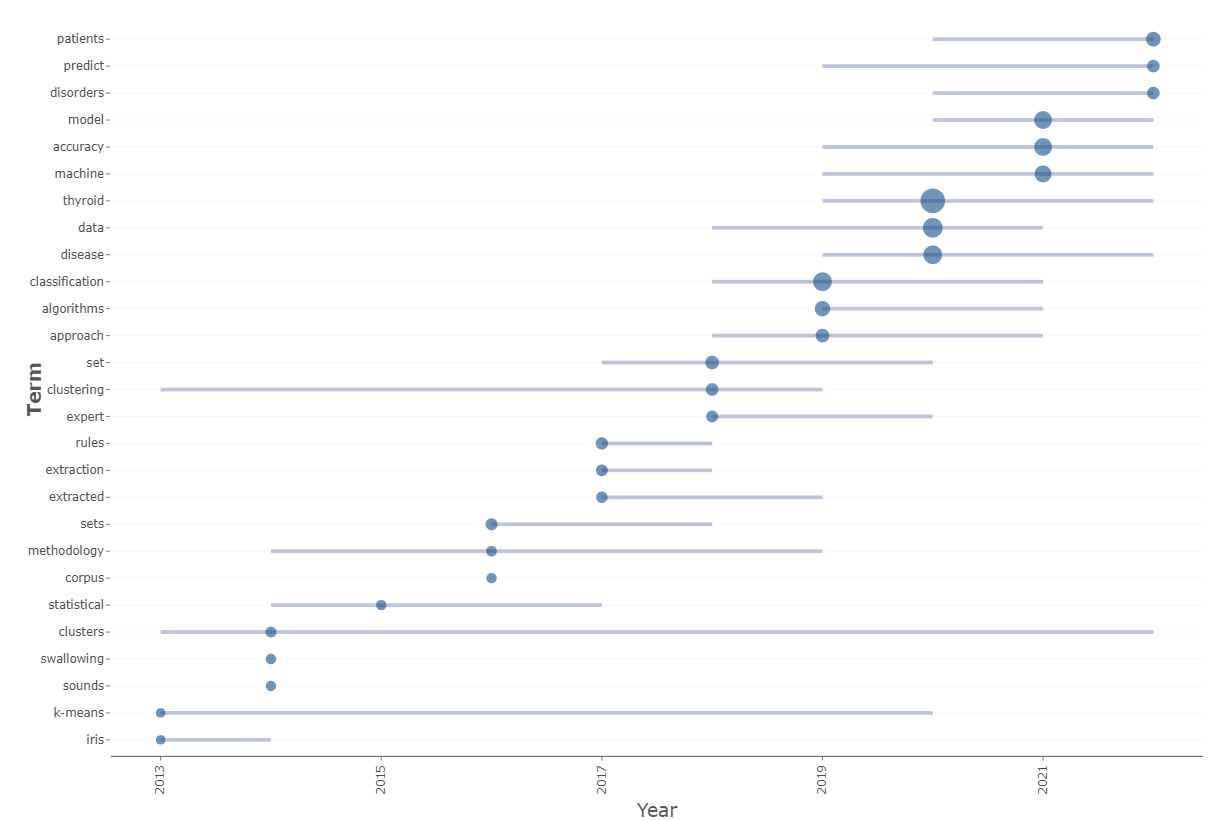




Fig. 7 Trending topics extracted from the topic of thyroid disease prediction

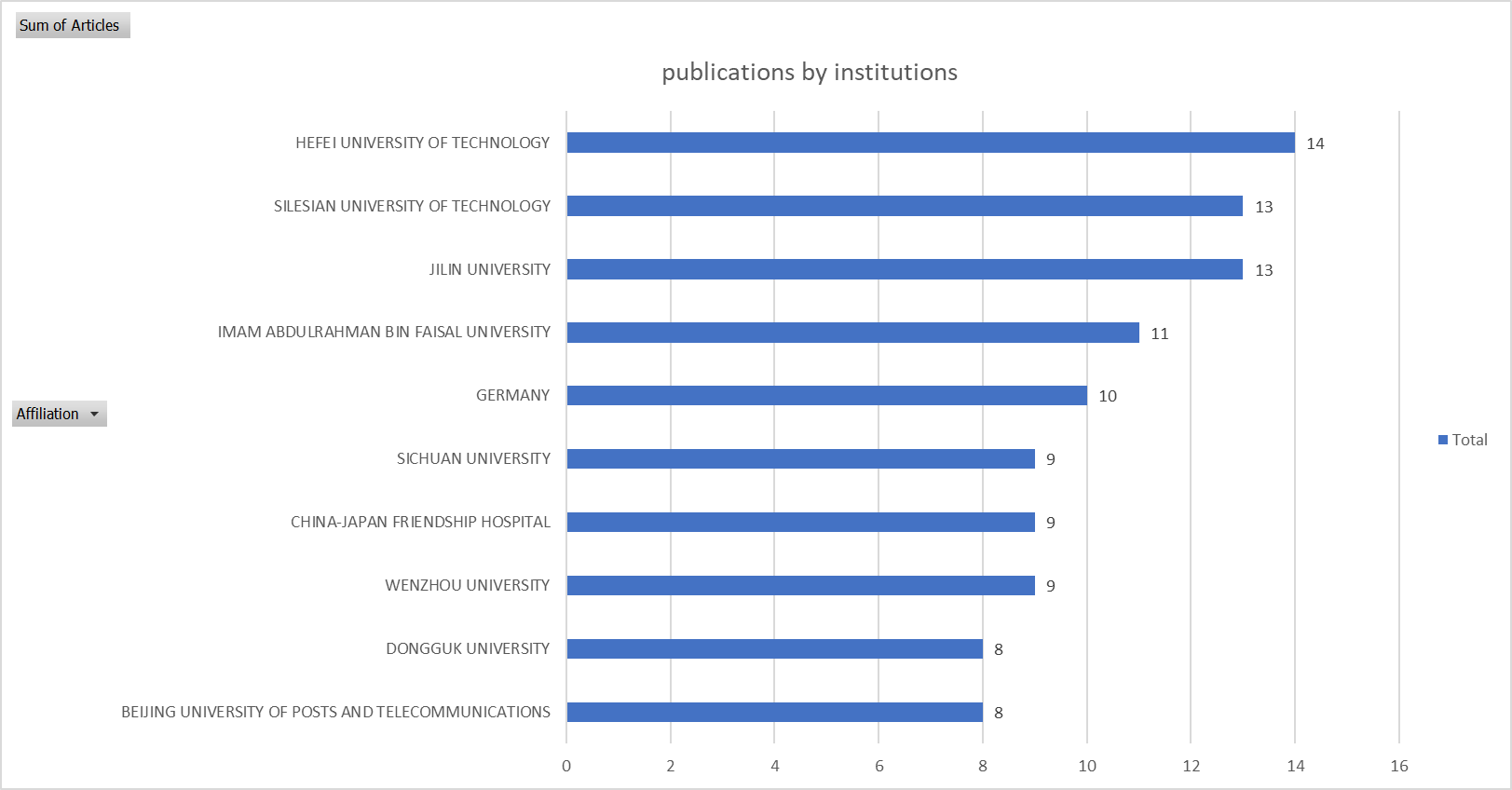
****

Fig. 8 top ten institutions based on the number of publications.

## Insights of TDDBML

In this section, we will look at 41 research articles that cover topics such as feature selection, unbalanced data, thyroid disease, and machine learning. It is intended that this in-depth examination of those 41 papers will shed light on the concept, methods, and potential future applications of interest to theorists and practitioners.

### Thyroid disease kinds

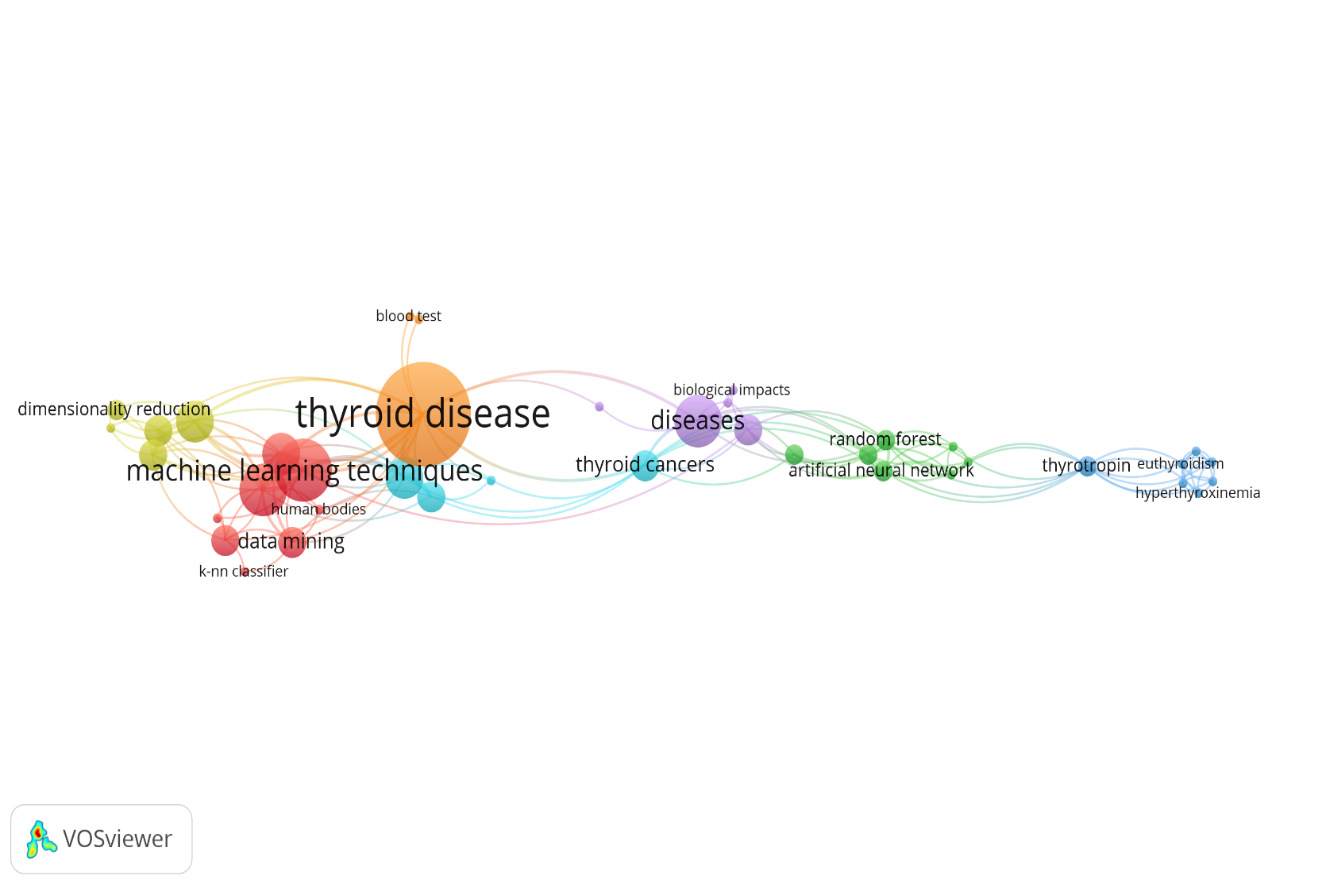
As machine learning-based methods improved, scientists and doctors began using data-driven methods to determine if a patient had a thyroid issue from a blood sample. Patients, however, often have to wait until their symptoms have worsened before they see a doctor because of the difficulties involved in undertaking the numerous routine tests. However, ML-based approaches enable early-stage diagnosis, which the subject himself can routinely perform using inexpensive and compact sensors [10]. Hypothyroidism, euthyroidism, goiter, thyroiditis, thyroid cancer, thyroid hormone resistance, and hyperthyroidism are seven diagnostic categories for thyroid disease. 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 conditions are severe and need medical attention.

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[22]. In order to effectively treat patients with thyroid disease, it is crucial to diagnose the condition early. For instance, Ahmed et al. achieved a 98.2% accuracy rate when training a deep neural network to differentiate between hypothyroid and hyperthyroid states[23]. Pal et al. compared the three machine learning models for predicting thyroid disease, including KNN, DT, and multilayer perceptron (MLP), and found that it achieves the highest accuracy of 94.23% [24]. 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 noncancer raw data using real-world data with 99% accuracy[25].

In machine learning techniques, most researchers use table-form data to detect multiple thyroid cancer. Instead of defining the disease, most researchers used the general term thyroid disease. Examples of studies that have reported results when accounting for thyroid disease by using a thyroid dataset include those by [24][26]. There are no specific thyroid diseases in the dataset, but rather 30 attributes that can be used to determine whether or not a patient has thyroid disease.

Fig. 9 depicts the most commonly reported disease associated with the thyroid from cited studies. Compared to other key-related diseases, "thyroid disease" has the largest cluster in the Figure; the two most common types of thyroid disease are also displayed, hyperthyroidism and hypothyroidism. 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.

Fig. 9 Illustration of most-reported thyroid disease (developed by VOSviewer software).



### Machine learning algorithms

TABLE IV illustrates that support vector machine (SVM) algorithms have received more attention from researchers and practitioners than any other ML type in designing PTDBML models over the years. 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 common technique in healthcare system prediction[29]. For instance, Płuciennik et al. have developed a model for thyroid cancer diagnostics which achieved approximately 95% accuracy[10]. Vairale et al. compared SVM to Logistic Regression (LR), K-NN, and NN for identifying people with the hypothyroid disease on the actual case dataset. SVM showed the best performance among all algorithms, producing an accuracy of 99%[30].

On the Other side, the RF classifier is the following algorithm to enhance the thyroid disease prediction model, which was nine studies conducted to develop a model for thyroid disease prediction. Alghamdi has designed an efficient predictive model to find thyroid cancer in the  Prostate, Lung, Colorectal and Ovarian (PLCO) dataset, defined as 155000 examples[31]. 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 PTDBML model development efforts have focused on DL algorithms rather than classic ML. Only 8 out of 41 studies focused on using DL to create a model for PTDBML, 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% using only 7 or 11 features [32]. M Asif et al. proposed that a Multilayer Perceptron (MLPC) was the most effective algorithm, achieving an accuracy of 99.70% after improving its hyperparameters [33]. In addition, Zhou et al. used ten ML algorithms through thyroid surgery to demonstrate a corresponding model. A CNN model can use Auc and accuracy to identify patients at an early stage of thyroid disease. Their main finding is that using data from 500 actual patients, the model can accurately identify those with thyroid disease 90% of the time and with 83% Auc. Other ML-based algorithms used by researchers to create the PTDBML model include KNN [9][11][31], Hoeffding [35], XGBoost [22], and Adaboost and Bagging [36].

### Imbalance challenges

One of the initial focuses was to track down previous research publications on the field of 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. Recent studies have addressed the issue of imbalanced data's effect on model performances, which most of the 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[37]. 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[31]. S. Aljameel et al. worked on a dataset that had an imbalance (much more thyroid cancer cases than nonthyroid cancer cases); thus, they utilized the SMOTEENN technique to avoid biasing the models toward one of the outcomes[25]. While [9], [38], [39] rely on SMOTE to handle imbalanced data issues to prevent bias in the performance measures. In cases where SMOTE is used to equalize the data, the overall model accuracy increases. 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[40].

Some researchers adopt DL-based solutions to replace all other algorithm-level methods. For instance, Selwal & Raoof developed a more accurate thyroid disease prediction system was used the MLP machine learning model and tested it on random samples that included hyperthyroid, hypothyroid, and healthy subjects[32]. 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 AUC 72 may be achieved using their proposed model. Thyroid disease-related datasets are notoriously unbalanced, and few publications have investigated methods to address this issue outside of classification and fabricated models[11], [16], [41]. Many ways are available to handle imbalanced classes; however, few studies have been mentioned in this overview that impact model performance the most.

### High dimensional challenges

Anotherinitial focus was to track down previous research publications on the thyroid disease field that included high dimensional data analyses. There are some studies which rely on optimizing features for datasets. In this way, the quality assessment revealed that 27 papers reported experimental results that originated from the feature selection or extraction dataset. For instance, Alghamdi has utilized Gini Algorithm(GA) to compute feature significance, and the model achieved an accuracy of 100% with both algorithms, RF and GB [31]**.** Some models have conducted a principal component analysis (PCA) to decrease the feature unrelated to the label [31][39][41].Arjaria et al. used SHAP (SHapley Additive explanations) strategies for values of each feature to reflect the signed influence of each characteristic on overall system performance [41]. Akhtar et al. designed a model to produce an efficient homogeneous ensemble of various feature-selection approaches to enhance thyroid illness identification, such as the Select K-Best (SKB), Select From Model (SFM), and Recursive Feature Elimination (RFE) [45]**.**

The authors [12], [27], [35], [46] defined feature selection as deleting irrelevant and redundant features from data collection to enhance the performance of machine learning techniques and their applications. In our reviewed studies, some studies used WEKA combining the feature selection strategy with the classification methodology[47], [48]. However, other studies did not conduct any techniques for decreasing or selecting features, leading to inaccurate diagnoses and critical errors that may lead to incorrect treatment and unnecessary costs for the right person due to the increasing number of high-dimensional features and data [49].

# **Discussions**

A thorough examination of 41 studies was done to understand the current practices and techniques used for identifying thyroid disorders when working with a high-dimensional and unbalanced dataset. The comprehensive analysis evaluated the following factors: thyroid disease type, applications, machine learning (ML) algorithms, high-dimensional, and imbalance solutions.

Overall, hypothyroidism and hyperthyroidism have received excellent attention in PTDBML. At the same time, other investigations looked into euthyroid [9], thyroid surgery[37], and so on. Sick-euthyroid syndrome is one of the most common problems in people suffering from malnutrition, trauma, surgery, or severe acute or chronic disease. The Due to a large number of features and little training data, the rates of the algorithms varied in this instance on cost-based feature selection[50]. Consequently, researchers and practitioners may focus on all types of thyroid disease rather than hypothyroidism and hyperthyroidism. Most ML-based models are designed to detect thyroid disease patients, emphasizing feature selection, and 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[30], [37], [51] and large datasets[32], [41]. A large amount of data helps the healthcare industry create more effective disease detection and decision-support systems[29]. The performance variation of the model [[1]](#endnote-1)is 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. In order to evaluate the effectiveness of ML-based models, it is required to use actual data rather than public repository data.

The instability of the model is one of the primary factors contributing to the restricted capabilities of the Clinical Decision Support Systems (CDSS) system. Since clinical systems cannot function correctly using only old patient data, the CDSS model must be continuously refined and updated, considering new information. 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.

Several machine learning (ML) methods, including SVM, RF, DT, KNN, ANN, MLP, and NB, are employed to construct the PTDBML model. However, the RF-based model attracts the most excellent attention from researchers since it provides the most accurate classification. Thousands of variables can be accommodated by random forest. Moreover, SVM is more successful in high dimensional areas and is relatively memory efficient, two reasons why many literature reviews have lately reported it. In situations when there are more dimensions than samples, SVM excels. Using data level, the problems of data ratio imbalance have been mainly addressed. SMOTE remains the most popular data-level solution, as seen by citations from only a few years ago. Unbalanced class solutions, which are more common in recent publications, have also contributed to the popularity of the DL-based method among academics. One of the problems with the DL-based method is that it does not explain how the model arrives at its conclusions.

Machine learning models often perform better intra-patient than inter-patient (inter-patient). Different data or patient characteristics can cause this. 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 another dataset. This could be due to patient-specific data, such as symptoms or blood test findings. To increase a machine learning model's performance on inter-patient data, a larger diverse dataset of patients may be needed to train the model to generalize to a broader range of patient populations. Adjusting the model's parameters or choosing a different machine-learning method may be essential.

One area of theoretical development has been using feature engineering to improve the predictive power of machine learning models. Feature engineering involves identifying and selecting the dataset's most relevant and predictive features variables and using these features to train the model[52]. This can help reduce the model's complexity and improve its generalizability to new data. It is also challenging to irrelevant features distinguish thyroid diseases from relevant features[39]. Consequently, although many researchers stated that their suggested model could perform well with high dimensional data, the model would only be trusted if its results were presented with an emphasis on feature selection, dimensionality reduction, regularization, and explainable ensemble approaches.

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[53], 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[44]. Implementing such a model in the actual world would be difficult and might be a fascinating area for future research. This is particularly relevant when the consequences of misclassification are severe, such as in 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 [54]. 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. The model's accuracy may vary depending on factors such as the location's geology, the size of the data set used, and the different types of thyroid disease being modelled. Some ML models, like those used to identify hypothyroidism, may not apply to other conditions. Since the training data for each disease is likely to come from a different source, developing an ML-based clinical detection method may require several models. Because of this, they were developing a framework that could deliver various disease determinations in real time would have to be complicated. The stability of the model through variable parameters or model updates based on user experience is also crucial to a safe diagnosis.

Table IV collects the results of the 41 studies cited in the literature to provide light on the ML-based prediction of thyroid disease.

# **Conclusions**

This study seeks to discover the most recent ML-based and data-driven developments and strategies in diagnosing thyroid disease using imbalanced and high-dimensional data. To develop ML-based systems for predicting thyroid disease in the real world, it is essential to enhance the ML-based experiments to include real-data patients and interpretable machine learning to explain the final prediction adequately. A comprehensive review of 41 papers suggests that more research is needed to prove reliable performance in healthcare settings. However, Deep Learning has come to dominate the area. Many academics and practitioners still use SMOTE as an Over-Sampling technique for handling unbalanced data.

Furthermore, PCA was still employed as a preprocessing stage for machine learning techniques owing to its ability to simplify the data by eliminating extraneous features and noise.

| **Author** | **Title** | **Algorithms for classification** | **Feature optimization** | **Imbalance** | **Evaluation** | **Dataset** |
| --- | --- | --- | --- | --- | --- | --- |
| [10] | Data Integration–Possibilities of Molecular and Clinical Data Fusion on the Example of Thyroid Cancer Diagnostics | SVM classifier (linear kernel) | Wilcoxon,  Relief | NA | p value below 0.05 | 200 real case |
| [37] | 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 | XGB, LGBM, and GBDT | ROCNA AUC curves | 92,91,91,91,91,90, 89,90, 89, 90 | 500 real case |
| [51] | Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset | RF, SVM, ANN, and NB | correlation coefficient | NA | 90, 84, 88, 81 | 218 real case |
| [31] | Evaluation of classification models for predicting mortality rate using thyroid cancer data | LR, KN, SVC, GNB,DT, AdaB ,RF and GB | Gini algorithm | Under-sampling | 98, 98, 99,  99,99, 99, 1, 1 | 155000 |
| [25] | A Proactive Explainable Artificial Neural Network Model for the Early Diagnosis of Thyroid Cancer | explainable artificial neural network (EANN) | remove | SMOTEENN | 0.98 | 724 real case |
| [9] | Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study | ANN, CatBoost, XGB, RF, LGBM, DT, SVC, KNN ,GNB | PCA | SMOTE | 95, 95, 95, 94,94, 94, 91, 89, 86 | 3,162  UCI |
| [22] | Thyroid Disease Prediction Using XGBoost Algorithms | LR, DT, KNN, and XGB | XGB | NA | 81, 87, 96, 98 | 215 UCI |
| [45] | thyroid disorder Effective voting ensemble of homogenous ensembling with multiple attribute selection approaches for improved identification | DT, GB,LR, RF | SFM, SKB, RFE | NA | 1 | 309 UCI |
| [40] | Use of the recursive rule extraction algorithm with continuous attributes to improve diagnostic accuracy in thyroid disease | Re RX | NA | BRACID | 96.70 | 7200 UCI |
| [55] | Butterfly Optimized Feature Selection with Fuzzy CNA Means Classifier for Thyroid Prediction | fuzzy CNA means algorithm (FCM) | DENA BOA | NA | 0.943 | 4152  UCI |
| [56] | Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification | ELM | fuzzy adaptive feature filtration | NA | 99.68 | 12944  UCI |
| [41] | Developing an Explainable Machine Learning Based Thyroid Disease Prediction Model | LR | SHAP | NA | 91 | 215  UCI |
| [34] | Increasing the Prediction Accuracy for Thyroid Disease: A Step Towards Better Health for Society | KNN  NN | PCA,SVD, DT | NA | 94, 98 | 3152  UCI |
| [32] | A Multi- layer perceptron-based intelligent thyroid disease prediction system | MLP | remove | NA | 99 | 120  UCI |
| [12] | Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning | DT, ANN | remove redundant features | NA | 99, 99 | 3772  UCI |
| [11] | Constituent depletion and divination of hypothyroid prevalance using machine learning classification | RF, DT, NB, KNN, LR | Ada Boost,  PCA | NA | 99 | 3164  UCI |
| [47] | Constructing a system for analysis of machine learning techniques for early detection of thyroid | ZeroR J48 Naïve bayes OneR | WEKA | NA | 60, 68, 41, 64 | 1,000  UCI |
| [35] | Decision tree ensemble techniques to predict thyroid disease | J48, RT, Hoeffding | remove | NA | 99, 97, 92 | 499  UCI |
| [44] | Prediction of Thyroid isease(Hypothyroid) in Early Stage Using Feature Selection and Classification Techniques | SVM, DT, RF, LR, NB | RFE, UFS, PCA | NA | 99, 99, 99, 96 | 519  UCI |
| [38] | Thy- Sys: A Preliminary Thyroid Wellness Assessment Through Machine Learning Using Pathological Factors | SVM, KNN, DT, SVMNA KNN | pathological features | SMOTE | 99.5 | 1464  UCI |
| [48] | Predictive Analysis for Thyroid Diseases Diagnosis Using Machine Learning | KNN, NB, DT | WEKA | NA | 92, 95, 99 | 1464 UCI |
| [33] | Computer-aided diagnosis of thyroid disease using machine learning algorithms | KNN, SVM, AdB, XGB, GPC, GBC,MLPC | Correlation | NA | 93, 96, 97, 96, 95, 98, 99 | 3164  UCI |
| [27] | A Machine Learning Approach to Predict Thyroid Disease at Early Stages of Diagnosis | DT, NB | remove | NA | 95 | 3000  UCI |
| [43] | Prediction of thyroid disorders using advanced machine learning techniques | NB, SVM, RF | RFE, US, TBFS, PCA | NA | 74, 92, 78 | 7200  UCI |
| [30] | Classification of Hypothyroid Disorder using Optimized SVM Method | KNN, SVM, LR, NN | remove | NA | 97, 99, 95, 94 | 574 real case |
| [46] | A Study on Label TSH, T3, T4U, TT4, FTI in Hyperthyroidism and Hypothyroidism using Machine Learning Techniques | RF, SVM, KNN | remove | NA | 98, 97, 95 | 7200  UCI |
| [42] | Feature selection algorithms to improve thyroid disease diagnosis | MLP BPNN SVM ELM | FNA Score, RFE, PCA | NA | 94, 95, 97, 98 | 215  UCI |
| [24] | Enhanced Prediction of Thyroid Disease Using Machine Learning Method | KNN, DT, MLP | Correlation | NA | 91, 94, 96 | 3163  UCI |
| [39] | Accuracy Assessment of Machine Learning Algorithm(s) in Thyroid Dysfunction Diagnosis | J48, MLP, NB, RF, SVM | NA | SMOTE | 99, 98, 98, 99, 98 | 4975  UCI |
| [36] | Efficient Thyroid Disease Prediction using Features Selection and Meta-Classifiers | Ada Boosting Bagging | NA | NA | 93, 99 | 774 |

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, we predict that ML-based thyroid disease detection with imbalanced data, dimensionality reduction, and other novel approaches may reveal numerous unrealized possibilities in the near future.

**REFERENCES**

[1] M. Clinic, “Hypothyroidism (underactive thyroid),” 2021, [Online]. Available: https://www.mayoclinic.org/diseases-conditions/hypothyroidism/symptoms-causes/syc-20350284

[2] World Health Organization (WHO), “pepole death by Thyroid diseases,” 2020.

[3] American Cancer Society, “Thyroid Cancer,” 2021. https://www.cancer.org/cancer/thyroid-cancer.html

[4] M. Obschonka and D. B. Audretsch, “Artificial intelligence and big data in entrepreneurship: a new era has begun,” *Small Bus. Econ.*, vol. 55, no. 3, pp. 529–539, 2020, doi: 10.1007/s11187-019-00202-4.

[5] V. Suteau, M. Munier, C. Briet, and P. Rodien, “Sex Bias in Differentiated Thyroid Cancer,” *Int. J. Mol. Sci.*, vol. 22, no. 23, p. 12992, 2021.

[6] M.-J. Schlumberger, S. Filetti, E. K. Alexander, and I. D. Hay, “Nontoxic diffuse goiter, nodular thyroid disorders, and thyroid malignancies,” in *Williams Textbook of Endocrinology*, Elsevier, 2016, pp. 449–488.

[7] M. Garbuio and N. Lin, “Artificial intelligence as a growth engine for health care startups: Emerging business models,” *Calif. Manage. Rev.*, vol. 61, no. 2, pp. 59–83, 2019, doi: 10.1177/0008125618811931.

[8] M. Alfifi, M. S. Alrahhal, S. Bataineh, and M. Mezher, “Enhanced artificial intelligence system for diagnosing and predicting breast cancer using deep learning,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 498–513, 2020, doi: 10.14569/IJACSA.2020.0110763.

[9] S. S. Islam, M. S. Haque, M. S. U. Miah, T. Bin Sarwar, and R. Nugraha, “Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study,” *PeerJ Comput. Sci.*, vol. 8, pp. 1–35, 2022, doi: 10.7717/PEERJ-CS.898.

[10] A. Płuciennik, A. Płaczek, A. Wilk, S. Student, M. Oczko-Wojciechowska, and K. Fujarewicz, “Data Integration–Possibilities of Molecular and Clinical Data Fusion on the Example of Thyroid Cancer Diagnostics,” *Int. J. Mol. Sci.*, vol. 23, no. 19, 2022, doi: 10.3390/ijms231911880.

[11] M. Shyamala Devi, A. Shil, P. Katyayan, and T. Surana, “Constituent depletion and divination of hypothyroid prevalance using machine learning classification,” *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 1607–1612, 2019, doi: 10.35940/ijitee.L3150.1081219.

[12] K. Guleria, S. Sharma, S. Kumar, and S. Tiwari, “Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning,” *Meas. Sensors*, vol. 24, no. September, p. 100482, 2022, doi: 10.1016/j.measen.2022.100482.

[13] A. J. Firestone and J. Settleman, “A three-drug combination to treat BRAF-mutant cancers,” *Nat. Med.*, vol. 23, no. 8, pp. 913–914, 2017.

[14] M. M. Ahsan *et al.*, “Detecting SARS-CoV-2 from chest X-Ray using artificial intelligence,” *Ieee Access*, vol. 9, pp. 35501–35513, 2021.

[15] W. Feng, W. Huang, and J. Ren, “Class imbalance ensemble learning based on the margin theory,” *Appl. Sci.*, vol. 8, no. 5, p. 815, 2018.

[16] K. S. Lee and H. Park, “Machine learning on thyroid disease: a review,” *Front. Biosci. - Landmark*, vol. 27, no. 3, 2022, doi: 10.31083/j.fbl2703101.

[17] A. M. Mendoza and R. M. Hernandez, “Application of Data Mining Techniques in Diagnosing Various Thyroid Ailments: A Review,” *Proc. 2021 13th Int. Conf. Inf. Commun. Technol. Syst. ICTS 2021*, pp. 207–212, 2021, doi: 10.1109/ICTS52701.2021.9608400.

[18] S. Anari, N. Tataei Sarshar, N. Mahjoori, S. Dorosti, and A. Rezaie, “Review of Deep Learning Approaches for Thyroid Cancer Diagnosis,” *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/5052435.

[19] C. Okoli and K. Schabram, “Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research,” *Work. Pap. Inf. Syst.*, vol. 10, no. 2010, 2010, doi: 10.2139/ssrn.1954824.

[20] B. Fahimnia, J. Sarkis, and H. Davarzani, “Green supply chain management: A review and bibliometric analysis,” *Int. J. Prod. Econ.*, vol. 162, pp. 101–114, 2015, doi: 10.1016/j.ijpe.2015.01.003.

[21] C. L. Borgman, “Big Data, Little Data, No Data: Scholarship in the Networked World—Who is in Charge of Data Quality.” MIT Press, 2015.

[22] S. Sankar, A. Potti, G. Naga Chandrika, and S. Ramasubbareddy, “Thyroid Disease Prediction Using XGBoost Algorithms,” *J. Mob. Multimed.*, vol. 18, no. 3, pp. 917–934, 2022, doi: 10.13052/jmm1550-4646.18322.

[23] I. Ahmed, R. Mohiuddin, M. A. Muqeet, J. A. Kumar, and A. Thaniserikaran, “Thyroid Cancer Detection using Deep Neural Network,” *Proc. - Int. Conf. Appl. Artif. Intell. Comput. ICAAIC 2022*, no. Icaaic, pp. 166–169, 2022, doi: 10.1109/ICAAIC53929.2022.9792854.

[24] M. Pal, S. Parija, and G. Panda, “Enhanced Prediction of Thyroid Disease Using Machine Learning Method,” *Proc. IEEE VLSI DCS 2022 3rd IEEE Conf. VLSI Device, Circuit Syst.*, vol. 2022, no. February, pp. 199–204, 2022, doi: 10.1109/VLSIDCS53788.2022.9811472.

[25] S. S. Aljameel, “A Proactive Explainable Artificial Neural Network Model for the Early Diagnosis of Thyroid Cancer,” *Computation*, vol. 10, no. 10, 2022, doi: 10.3390/computation10100183.

[26] A. Tyagi, R. Mehra, and A. Saxena, “Interactive thyroid disease prediction system using machine learning technique,” *PDGC 2018 - 2018 5th Int. Conf. Parallel, Distrib. Grid Comput.*, pp. 689–693, 2018, doi: 10.1109/PDGC.2018.8745910.

[27] A. R. Rao and B. S. Renuka, “A Machine Learning Approach to Predict Thyroid Disease at Early Stages of Diagnosis,” *2020 IEEE Int. Conf. Innov. Technol. INOCON 2020*, pp. 2020–2023, 2020, doi: 10.1109/INOCON50539.2020.9298252.

[28] K. Salman and E. Sonuc, “Thyroid Disease Classification Using Machine Learning Algorithms,” *J. Phys. Conf. Ser.*, vol. 1963, no. 1, 2021, doi: 10.1088/1742-6596/1963/1/012140.

[29] V. Kamra, P. Kumar, and M. Mohammadian, “Diagnosis support system for general diseases by implementing a novel machine learning based classifier,” *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 737–746, 2021, doi: 10.12785/ijcds/100168.

[30] V. S. Vairale and S. Shukla, “Classification of Hypothyroid Disorder using Optimized SVM Method,” *Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019*, no. Icssit, pp. 258–263, 2019, doi: 10.1109/ICSSIT46314.2019.8987767.

[31] N. S. Alghamdi, “Evaluation of classification models for predicting mortality rate using thyroid cancer data,” *J. Comput. Sci.*, vol. 15, no. 1, pp. 131–142, 2019, doi: 10.3844/jcssp.2019.131.142.

[32] A. Selwal and I. Raoof, “A Multi-layer perceptron based intelligent thyroid disease prediction system,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 17, no. 1, pp. 524–532, 2020, doi: 10.11591/ijeecs.v17.i1.pp524-532.

[33] M. A. A. R. Asif *et al.*, “Computer aided diagnosis of thyroid disease using machine learning algorithms,” *Proc. 2020 11th Int. Conf. Electr. Comput. Eng. ICECE 2020*, vol. 4, pp. 222–225, 2020, doi: 10.1109/ICECE51571.2020.9393054.

[34] R. Jha, V. Bhattacharjee, and A. Mustafi, “Increasing the Prediction Accuracy for Thyroid Disease: A Step Towards Better Health for Society,” *Wirel. Pers. Commun.*, vol. 122, no. 2, pp. 1921–1938, 2022, doi: 10.1007/s11277-021-08974-3.

[35] D. C. Yadav and S. Pal, “Decision tree ensemble techniques to predict thyroid disease,” *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 8242–8246, 2019, doi: 10.35940/ijrte.C6727.098319.

[36] D. Priyadharsini and S. Sasikala, “Efficient Thyroid Disease Prediction using Features Selection and Meta-Classifiers,” *Proc. - 6th Int. Conf. Comput. Methodol. Commun. ICCMC 2022*, no. Iccmc, pp. 1236–1243, 2022, doi: 10.1109/ICCMC53470.2022.9753986.

[37] C. M. Zhou, Y. Wang, Q. Xue, J. J. Yang, and Y. Zhu, “Predicting difficult airway intubation in thyroid surgery using multiple machine learning and deep learning algorithms,” *Front. Public Heal.*, vol. 10, 2022, doi: 10.3389/fpubh.2022.937471.

[38] I. R. Francisco, M. B. J. Ferolin, C. F. Pena, and R. J. Ferolin, “Thy-Sys: A Preliminary Thyroid Wellness Assessment Through Machine Learning Using Pathological Factors,” *Proc. - 2021 1st Int. Conf. Inf. Comput. Res. iCORE 2021*, pp. 44–49, 2021, doi: 10.1109/iCORE54267.2021.00027.

[39] K. J. Danjuma, G. Maksha Wajiga, E. J. Garba, A. Sandra Ahmadu, and O. B. Longe, “Accuracy Assessment of Machine Learning Algorithm(s) in Thyroid Dysfunction Diagnosis,” *Proc. 2022 IEEE Niger. 4th Int. Conf. Disruptive Technol. Sustain. Dev. NIGERCON 2022*, 2022, doi: 10.1109/NIGERCON54645.2022.9803113.

[40] Y. Hayashi, S. Nakano, and S. Fujisawa, “Use of the recursive-rule extraction algorithm with continuous attributes to improve diagnostic accuracy in thyroid disease,” *Informatics Med. Unlocked*, vol. 1, no. 2015, pp. 1–8, 2015, doi: 10.1016/j.imu.2015.12.003.

[41] S. K. Arjaria, A. S. Rathore, and G. Chaubey, “Developing an Explainable Machine Learning-Based Thyroid Disease Prediction Model,” *Int. J. Bus. Anal.*, vol. 9, no. 3, pp. 1–18, 2022, doi: 10.4018/IJBAN.292058.

[42] K. Pavya and B. Srinivasan, “Feature selection algorithms to improve thyroid disease diagnosis,” *IEEE Int. Conf. Innov. Green Energy Healthc. Technol. - 2017, IGEHT 2017*, pp. 1–5, 2017, doi: 10.1109/IGEHT.2017.8094070.

[43] P. Duggal and S. Shukla, “Prediction of thyroid disorders using advanced machine learning techniques,” *Proc. Conflu. 2020 - 10th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 670–675, 2020, doi: 10.1109/Confluence47617.2020.9058102.

[44] M. Riajuliislam, K. Z. Rahim, and A. Mahmud, “Prediction of Thyroid Disease(Hypothyroid) in Early Stage Using Feature Selection and Classification Techniques,” *2021 Int. Conf. Inf. Commun. Technol. Sustain. Dev. ICICT4SD 2021 - Proc.*, pp. 60–64, 2021, doi: 10.1109/ICICT4SD50815.2021.9397052.

[45] T. Akhtar *et al.*, “Effective voting ensemble of homogenous ensembling with multiple attribute-selection approaches for improved identification of thyroid disorder,” *Electron.*, vol. 10, no. 23, 2021, doi: 10.3390/electronics10233026.

[46] A. H. Shahid, M. P. Singh, R. K. Raj, R. Suman, D. Jawaid, and M. Alam, “A Study on Label TSH, T3, T4U, TT4, FTI in Hyperthyroidism and Hypothyroidism using Machine Learning Techniques,” *Proc. 4th Int. Conf. Commun. Electron. Syst. ICCES 2019*, no. Icces, pp. 930–933, 2019, doi: 10.1109/ICCES45898.2019.9002284.

[47] S. Rasheeduddin and K. Rajasekhar Rao, “Constructing a system for analysis of machine learning techniques for early detection of thyroid,” *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6 Special Issue 3, pp. 1978–1981, 2019, doi: 10.35940/ijeat.F1385.0986S319.

[48] Z. J. Peya, M. K. N. Chumki, and K. M. Zaman, “Predictive Analysis for Thyroid Diseases Diagnosis Using Machine Learning,” *2021 Int. Conf. Sci. Contemp. Technol. ICSCT 2021*, pp. 4–9, 2021, doi: 10.1109/ICSCT53883.2021.9642544.

[49] D. A. A. Gnana, “Literature Review on Feature Selection Methods for High-Dimensional Data Literature Review on Feature Selection Methods for High-Dimensional Data,” no. February, 2016, doi: 10.5120/ijca2016908317.

[50] Y. Zhang, D. W. Gong, and J. Cheng, “Multi-objective particle swarm optimization approach for cost-based feature selection in classification,” *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 14, no. 1, pp. 64–75, 2017, doi: 10.1109/TCBB.2015.2476796.

[51] S. O. Olatunji *et al.*, “Early diagnosis of thyroid cancer diseases using computational intelligence techniques: A case study of a Saudi Arabian dataset,” *Comput. Biol. Med.*, vol. 131, no. February, p. 104267, 2021, doi: 10.1016/j.compbiomed.2021.104267.

[52] G. Dong and H. Liu, *Feature engineering for machine learning and data analytics*. CRC Press, 2018.

[53] D. Gan, J. Shen, B. An, M. Xu, and N. Liu, “Integrating TANBN with cost sensitive classification algorithm for imbalanced data in medical diagnosis,” *Comput. Ind. Eng.*, vol. 140, no. June 2019, p. 106266, 2020, doi: 10.1016/j.cie.2019.106266.

[54] B. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering,” 2007.

[55] S. J. K. Jagadeesh Kumar, P. Parthasarathi, M. Masud, J. F. Al-Amri, and M. Abouhawwash, “Butterfly Optimized Feature Selection with Fuzzy C-Means Classifier for Thyroid Prediction,” *Intell. Autom. Soft Comput.*, vol. 35, no. 3, pp. 2909–2924, 2023, doi: 10.32604/iasc.2023.030335.

[56] K. Juneja, *Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification*, no. 0123456789. Springer US, 2022. doi: 10.1007/s11277-022-09823-7.

**Dhekre Saber Saleh** received a B.Sc. degree in Computer Science from the University Soran (SU), in the Kurdistan Region of Iraq in 2010 and the Master of degree in Computer Engineering from the Cyprus International University (CIU) in Northern Cyprus in 2015 and his PhD degree in Computing from the University Technology of Malaysia (UTM). His research interests include machine learning, optimization, big data, business intelligence, and data mining.

**Mohd Shahizan Othman** received a B.Sc. degree in computer science with a major in information systems from the Universiti Teknologi Malaysia (UTM), Malaysia, in 1998 and an M.Sc. degree in information technology and the Ph.D. degree in web information extraction, information retrieval and machine learning from the Universiti Kebangsaan Malaysia (UKM), Malaysia. He is currently an Associate Professor with the Faculty of Engineering, School of Computing, UTM. His research interests include information extraction and information retrieval on the web, web data mining, content management, machine learning, social learning, e-learning, business intelligence, and geographic information system (GIS).

**A person with a mustache and a suit and tie

Description automatically generated with low confidenceQasim Hamakhurshid Hamamurad** is a doctoral student at the University of Technology Malaysia, Information System Department. He received his BS in Civil Engineering from the University of Salahadin, Iraq-Kurdistan in 1996 and his 1st MSc in Computer science and Communication Technology from the University Amiens Picardy Jules Verne (UPJV) France country in 2013, while 2nd MSc degree in Management Information Systems in 2012 at Lebanese French University (LFU) respectively, and his PhD degree in Information System from the University Technology of Malaysia (UTM). He is the author of more high-impact journal articles and more than 13 Books. His currently an international academic lecturer. ORCID iD: <https://orcid.org/0000-0002-2684-5007>

**Wshyar Omar Khudhur** received a B.Sc. degree in software engineering from Koya University in the Kurdistan Region of Iraq, in 2008 and a master of philosophy degree in Computer Engineering from the Cyprus International University (CIU) in Northern Cyprus in 2015, and he is currently head of Information Technology Department in Koya Technical Institute – Erbil Polytechnic University (EPU).

**Eman Aljabarti** was born in Dammam, Saudi Arabia, in 1984. She received a B.E. degree in computer science from the Imam Abdulrahman bin Faisal University, Dammam, in 2005, and an MSc degree in computer science from Universiti Teknologi Malaysia (UTM) Johor Bahru, Johor, Malaysia, in 2022. Her current research interests include artificial intelligence, data analysis, healthcare, the education field of cloud computing, the internet of things, and big data.

**Saud S Alshammari** received a B.S. degree in computer information systems from the University of the Incarnate Word, Texas, USA in 2015, and an MSc degree in computer science from Universiti Teknologi Malaysia (UTM) Johor Bahru, Johor, Malaysia, in 2022. He joined the University of Hafer Albatin, Saudi Arabia, in 2018. His main areas of research interest are machine learning, data mining, big data, computer vision, computational linguistics & NLP.

1. [↑](#endnote-ref-1)