

أحمد عبدالله

Master Thesis Template-23-4-2026.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:379296215

Submission Date

Apr 23, 2026, 4:50 PM GMT+2

Download Date

Apr 24, 2026, 10:56 AM GMT+3

File Name

Master Thesis Template-23-4-2026.docx

File Size

92.9 KB

42 Pages

5,167 Words

32,199 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.





Hybrid Machine Learning Framework for Multi-Class Thyroid Risk Assessment

Prepared By

Name: Ahmed Abdullah

ID No: 253000598

Submitted in fulfillment of the requirements for the degree Master of
Science in Artificial Intelligence

At

Collage of Informatics

Midocean University

UAE

Supervisor Name: Dr. Amjad Qtaish

ii

للإستخدام الداخلي - Restricted Internal use

Collage of Informatics, Midocean University, Comoros

Email:.....

للإستخدام الداخلي - Restricted Internal use

Acknowledgements

Special thanks to Dr. .

Restricted Internal use - للإستخدام الداخلي

Restricted Internal use - للإستخدام الداخلي

Abstract

List of Publications

Table of Contents

List of Abbreviations	X
List of Figures.....	xi
List of Tables	xii

Chapter 1: Introduction and

Overview.....	1
1.2 Problem Definition	2
1.3 Thesis Objectives	3
1.4 Thesis Contribution	4
1.5 Thesis Organization.....	5

Chapter 2: Background and Related Work

.....	31
-------	----

Chapter 3: Proposed Approach 78

4.1 Introduction	78
------------------------	----

Chapter 4: Experimental Results and Discussion 86

Chapter 5: Conclusion and Future Work 127

7.1 Important Features	127
7.3 Future Work	129

References	130
------------------	-----

List of Abbreviations

- **Acc.:** Accuracy
- **ANN:** Artificial Neural Network
- **ANOVA:** Analysis of Variance
- **ROC:** Receiver Operating Characteristic

List of Figures

2.1 Educational Data Mining Process	7
2.2 Applications of Educational Data Mining	8
2.3 Distribution of Papers that used different attributes for PSP problem	19
2.4 Student Performance Data Set	21
2.5 Educational Process Mining (EPM): A Learning Analytics Data Set ...	23

List of Tables

2.1	Student Performance Data Set Description	21
2.2	Students' Academic Performance Dataset Description	23

Chapter 1 Introduction

1.1 Overview

Thyroid issues are one of the most common conditions worldwide and are known to affect a wide range of physiological processes within human beings. The thyroid gland has a critical direct impact on regulating metabolism, cardiovascular function, and body temperature. Any disturbance in thyroid hormone production may lead to significant complications, including hypothyroidism and hyperthyroidism. Due to the prevalence of these conditions and their impact on human health, accurate and timely diagnosis remains a critical clinical point (Taylor et al., 2022).

Usually, thyroid issues are evaluated through lab testing of markers, such as thyroid-stimulating hormone (TSH), triiodothyronine (T3), and thyroxine (T4). These numbers are common in providing insights in clinical practice (Biondi & Cappola, 2021). This lab metric does not always provide a direct insight, especially when numbers fall near diagnostic thresholds or when a lot of indicators must be taken into consideration at the same time.

Recent practices done by technology / medical members in illness identification and classification. These practices provide support in analyzing medical data and lab results. In our case in this paper, many studies introduced the benefits of implementing machine learning models in the identification of conditions in the clinics based on lab numbers (Rajkomar et al., 2022). Many

good results have been proven through adopting these techniques in terms of accuracy.

Based on the facts mentioned, there is a gap that needs to be filled to not only provide accurate classification of thyroid disorders using lab results but also present the results in a clear, integrated method that can support clinical decisions. In this study we will address this need through implementing a machine learning framework that focuses on lab results while stressing interpretability and usability.

1.2 Problem Definition

Machine learning methods have been applied to prediction thyroid disease; most studies have mainly focused on improving classification using lab biomarkers such as TSH, T3, and T4. While these techniques have proven effective results, their adoption in real clinical practices remains limited (Biondi & Cappola, 2021). Frequently, models are implemented and adopted based on metrics alone, without sufficient consideration of how the results can be integrated and implemented by medical practitioners.

In clinical practice, diagnostic decisions are based on metrics and understanding of how different metrics contribute to the final consideration. However, many machine learning frameworks operate as mysterious systems, where predictions are generated without detailed and clear explanation. The absence of transparency reduces trust in these kinds of frameworks and limits their implementation in medical environments. This can result in even highly accurate models not providing meaningful support for clinical practices.

Also, the absence of datasets that have accurate data impacted the adoption of more accurate methodologies. These issues are presented many times by authors in their papers in chapter 2 (Literature Review); new papers still present their work focusing only on lab testing results. Through this, a gap was found between the actual performance and technology in supporting medical people on their daily activities.

By looking into my statements above, I can provide an urge for a framework that does not only focus on providing medical people a classification of illnesses but also provides an enhancement of using the lab test results. The future framework is expected to provide a better understanding and provide medical teams structured results that match clinical results.

1.3 Thesis Objectives

I have a lot of goals that I look forward to introducing in this paper, but the most important goal that I would like to introduce in this paper is a machine learning technique for thyroid problems with UCI laboratory datasets.

The research will be guided by the following guidelines to help us meet our objective:

- Use the UCI Thyroid data set and regard it as the major data set for our study.
- Adopt decision trees and random forests as machine learning methods.
- Test our findings/outputs using numbers and make sure they are accurate.
- Create a framework that provides results to medical practitioners in a way that they understand.

- Reveal the limitations and challenges encountered during this study and provide suggestions on how to improve it in the future.

By doing these things, I would offer a framework that will provide the combination of high performance medical practitioners can rely on and interpretability for advanced machine learning.

1.4 Thesis Contribution

First, the study will pinpoint weaknesses in existing studies concerning the status of thyroid prediction at present. The above will be done by emphasizing the need to enhance not only the performance of the models but also their interpretability and practical implementation.

Decision trees and random forests will be used in the analysis of thyroid prediction by means of the UCI Thyroid dataset and precision, recall, and F1 scores. The above-mentioned conceptual model is proposed so that findings can be communicated in an easy-to-understand way for medical practitioners.

Ultimately, issues with laboratory metric-based models will be determined, along with recommendations for the future.

1.5 Thesis Organization

Chapter 1

Chapter 1 will have an introduction to the problem and background information of the problem, and objectives of the work will be discussed in this chapter.

Chapter 2

In chapter 2, a literature review will be done and comparisons of the results will be done using similar papers. In the last part of this chapter, comparison of the similar paper work will be done.

Chapter 3

Chapter 3 will discuss the methodology that includes selection of the dataset and selecting appropriate methodology to accomplish the objective.

Chapter 4

Chapter 4 will discuss the results obtained from experimental work along with discussions on it. It will be done through presentation of the results using model evaluation measures of different classification models.

Chapter 5

Conclusion will be presented in the last chapter of the thesis.

Chapter 2: Background and Related Work

Thyroid disorder is considered an important disease from the medical aspect because of its high prevalence among humans and impacts on the body processes, such as metabolic rate, heart functions, and temperature regulation. In order to determine the efficiency of functioning of the thyroid gland, there is a need to evaluate different biochemical indices and their complex interconnection with the symptoms. This makes the computer-based application quite necessary because this system will be capable of analyzing clinical indices. Machine learning techniques have been proven to be rather useful in contemporary medicine because they allow for detecting regularities in health records that can be used for diagnosing different ailments. For this reason, many studies have utilized machine learning techniques for predicting or classifying diseases on the basis of different laboratory indices. Therefore, applying medical records of thyroid disorders with the use of machine learning techniques is rather required.

2.1 Thyroid Gland and Thyroid Disorders

There exist many individuals globally who are experiencing health problems related to the thyroid. The thyroid gland is responsible for regulating various metabolic activities within the body. The thyroid affects a number of biological activities such as metabolism rates, cardiovascular actions, thermoregulation, and energy levels, among others. This implies that any form of dysfunction in the levels of thyroid hormone production may affect various organs and systems in the body (Biondi & Cappola, 2021).

The thyroid gland is in the front part of the neck and produces two primary hormones: thyroxine (T4) and triiodothyronine (T3). These hormones are responsible for adjusting metabolism. The secretion of thyroid hormones is controlled by the hypothalamic–pituitary–thyroid (HPT) axis. In this scenario, the hypothalamus releases thyrotropin-releasing hormone (TRH), which increases the pituitary gland to produce thyroid-stimulating hormone (TSH). TSH then increases the thyroid gland to release T3 and T4. The levels of these hormones are regulated through a feedback mechanism that helps maintain hormonal balance (Ross et al., 2016).

Thyroid dysfunction happens when the hormonal system does not function as usual. The most common forms of thyroid dysfunction are hypothyroidism and hyperthyroidism. Hypothyroidism happens when the thyroid gland produces not enough levels of thyroid hormones. Patients with hypothyroidism may experience symptoms such as fatigue, weight gain, cold intolerance, and dry skin. Hyperthyroidism, on the other hand, occurs when the thyroid gland produces excessive thyroid hormones and may lead to symptoms such as weight loss, palpitations, anxiety, and heat intolerance (Biondi & Cappola, 2021; Chiovato et al., 2023).

To identify thyroid disorders, medical practitioners involve lab results of several biochemical indicators. Measurements of thyroid-stimulating hormone (TSH), triiodothyronine (T3), and thyroxine (T4) are commonly used to evaluate thyroid function. In some cases, additional derived indicators such as the Free Thyroxine Index (FTI) are used to improve the outcome understanding. Because thyroid hormones operate within a feedback system, involving these biomarkers

requires evaluating their relationships rather than examining each measurement alone (Ross et al., 2016).

Symptoms, in addition to lab indicators, also play an important role in the evaluation process. Yet, there is a lack of specificity with respect to symptoms since there are many symptoms that are not unique to thyroid disorders. These include fatigue, changes in weight, feeling hot or cold, and palpitations, among others. This causes clinicians to rely on a combination of laboratory indicators and evaluations done in the clinics when evaluating thyroid disorders (Chiovato et al., 2023).

The integration of various biochemical markers along with symptoms in a patient can sometimes be difficult, especially when lab values come very close to the cut-off values for diagnosis or there is a lack of definition in the symptoms. This has driven the need for computer-based algorithms capable of analyzing multiple parameters at once. Recent developments in AI and machine learning have revealed promising results for the analysis of health care data and predicting diseases (Khan et al., 2024).

It is crucial to have an understanding of the medical background regarding the physiology of the thyroid, its biomarkers, and symptoms before discussing the application of computational methods in predicting thyroid diseases. The knowledge acquired forms the basis for analyzing machine learning patterns that help in assessing thyroid data.

2.2 Machine Learning in Healthcare

Machine learning technology has become the new trend in contemporary analysis software applications in the health care sector. It involves software programs that allow machines to identify patterns in data and make predictions or classifications without requiring substantial coding for each task performed by the algorithm. Machine learning allows an algorithm to process big data and establish correlations among different variables that cannot be easily established through conventional statistics (Esteva et al., 2021).

In the health care industry, machine learning technology has increasingly been applied as the new trend in diagnosing diseases, assessing risk factors, and assisting decision-making processes. Medical data has several complicated relationships among patient data, laboratory reports, physical observations, and historical data. Machine learning technologies have been applied in processing such data and identifying patterns to assist physicians in their diagnoses and predictions (Topol, 2023).

Machine learning algorithms application in health care includes designing clinical decision support systems. The latter should provide clinical feedback through evaluation of data. The intention here is not to replace medical expertise but to complement the process through the provision of feedback which helps the practitioners understand information and detect disease patterns. Machine learning algorithms have found their way in many areas of medicine ranging from prediction of cardiovascular conditions, cancers diagnosis, and endocrine conditions analysis (Khan et al., 2024).

Through the machine learning patterns, the lab indicators and symptoms can easily be integrated and evaluated together. Diagnosing any thyroid condition requires evaluating several lab indicators and symptoms at the same time. This might be quite hard under traditional approaches. To overcome such challenges, computational models could be used to evaluate combinations of lab results and symptoms to establish probabilities of thyroid conditions.

For this case study, the purpose of using the proposed algorithm is to assist in evaluating the risks of developing thyroid conditions through integration of clinical symptoms and lab indicators.

2.3 Machine Learning for Thyroid Disease Prediction

Nowadays, machine learning models have been adopted in healthcare domains to aid disease identification and classification. Using machine learning models in the field of healthcare allows one to analyze the data obtained during practical activities and discover any patterns which might be related to thyroid disease. The connection between lab results and symptoms in thyroid diseases is rather complicated, which makes machine learning essential.

A number of research works have used machine learning models for predicting thyroid diseases. Most of those works have used laboratory parameters, including thyroid stimulating hormone (TSH), triiodothyronine (T3), and thyroxine (T4) as predictor variables. Several supervised machine learning models, such as decision tree, support vector machine, random forest, and neural network, were suggested as ways to classify patients according to their thyroid diseases (normal thyroid function, hypothyroidism, and hyperthyroidism).

Recently, researchers have focused on improving prediction performance by comparing various machine learning models. Many machine learning algorithms have been examined using structured data such as the UCI Thyroid Disease Data Set. From the findings, it is evident that machine learning models can yield positive results in predicting thyroid disease (Khan et al., 2024).

While significant improvements have been made regarding the thyroid diseases prediction issue by machine learning algorithms, almost all research papers have considered only the results from laboratory tests without mentioning the signs of patients. When diagnosing thyroid disease in the context of medicine, physicians consider the lab test results as well as signs of patients.

In addition, a lot of literature talks about the preciseness of prediction without considering any other criteria, for instance, interpretability. Thus, if we talk about health issues, the machine learning algorithms should be precise and at the same time interpretable to make decisions based on their results.

Several disadvantages inherent to the existing approaches applied in assessment of thyroid data have been listed in this paper. Having considered the mentioned disadvantages, it was suggested to use the hybrid approach which combines symptom assessment and machine learning algorithms.

2.4 Comments of other researchers

Thyroid disease has been studied by many researchers using different methods. Both machine learning laboratory methods and clinical assessment methods have been employed to determine diagnosis and/or prognosis. Much current research focuses on laboratory biomarkers as the input to the predictive model.

Along these lines, Islam et al. (2022) found that machine learning models can be highly effective at predicting hypothyroidism from biological indicators such as TSH, T3 and T4 levels during hospital training sessions. Similar outcomes were achieved by Alshamrani et al. (2023) by applying several machine learning techniques to organized thyroid datasets to achieve accurate classification results. While results from such data are significant and reliable, more recent analyses by Shen et al. (2023) indicate that many methods heavily rely on organized numerical data and focus on improving the accuracy of future predictions.

Several researchers explored in depth the clinical aspects of thyroid diseases, emphasizing the difficulty of the diagnosis and the limitations of an entirely laboratory-based approach. Biondi and Cappola (2021) highlighted the many links of thyroid dysfunction, and its correct interpretation requires a thoughtful clinical approach. Chiovato et al. (2023) stressed that, in the real-life clinical practice, the diagnosis of thyroid diseases mainly relies on the correct interpretation of the laboratory data in the frame of the whole clinical case. There has been growing interest in how AI can be utilized within clinical practice as part of AI in healthcare. In their review of decision-support systems, Sutton et al. (2020) suggested that output should be clear, transparent and appropriate.

Effectiveness and integration into clinical workflows (Jiang et al. 2022) are also key considerations. However, as noted by Sendak et al. (2021), many machine learning models fail to translate into clinical practice due to user unfriendliness and lack of trustworthiness. In addition, it is generally agreed that a model which achieves high performance in prediction must also have good interpretability and reproducibility (Miotto et al. 2021, Rajkomar et al. 2022).

Thyroid diseases in the lab have been much better characterized in recent years. However, a major limitation of current research is the lack of focus on the interpretability and applicability of model results. Even though good predictions can be made, the results are not interpretable enough and therefore not useful enough for clinical decision support.

2.4.1 Analytical Discussion of Previous Research

It therefore becomes clear that existing literature offers a wealth of information but fails to fulfill the required criteria. Machine learning studies conducted in laboratories have shown remarkable prediction capabilities; nevertheless, their heavy reliance on numerical measures makes it harder to comprehend their clinical significance, as mentioned in various reviews, including Shen et al. (2023).

Unlike that, clinical studies like the one carried out by Biondi & Cappola (2021) and Chiovato et al. (2023) demonstrate how complicated it is to analyze thyroid disorders and the requirement for the interpretation of findings. Nonetheless, there are no computational methods suggested by previous studies for automatic classification.

Furthermore, the mentioned properties that an AI-based system needs to have for being considered successfully implemented in the medical environment include interpretability, reliability, and clinically oriented nature, according to the literature on clinical applications, such as Sutton et al. (2020), Sendak et al. (2021), and Jiang et al. (2022). However, although there is a lot of evidence for the need for the discussed attributes, most algorithms are oriented exclusively towards increasing the accuracy of their models, as was indicated, among others, by Rajkomar et al. (2022) and Miotto et al. (2021).

Taking into consideration the presented limitation, the current paperwork opts to use machine learning based on numerical laboratory values, focusing specifically on its interpretability.

The idea behind such a choice implies that the emphasis should not be on the predicted value itself but rather on explaining how the behavior of the created algorithm operates. According to this purpose, one may assume that the research will be valuable for the classification of the discussed health condition.

2.5 Research Gap

In this regard, it must be said that these articles show a shift in the field in their departure from diagnostic techniques and application of more sophisticated technology in practice, such as machine learning. At the same time, no matter what new developments may appear in this field, the fact still exists that physicians do not want to gather all data obtained during laboratory analysis and during clinical diagnosis into a single database.

It can be exemplified by the guidelines proposed by Ross et al. (2016). According to the recommendations provided by the authors, physicians should diagnose and treat patients with various conditions related to the functioning of the thyroid gland according to data obtained as a result of laboratory analysis of blood, i.e., TSH, T3, and T4. Recommendations are created on the basis of experience and acquired medical knowledge; hence, there is no need for any sort of data modeling and management in this context.

However, now consider the example of Dua and Graff (2019) and the thyroid data from the UCI Machine Learning Repository. This is an excellent resource for those who need to prepare their data, select features, and run ML models. However, the dataset is numeric only, meaning no records about patient symptoms or narratives of cases can be found there. According to Shen et al. (2023), such datasets overvalue laboratory parameters in the analysis and undervalue the importance of the whole clinical picture.

Sutton et al. (2020) concentrate on the usability of the systems being discussed from a wide range of perspectives. Specifically, the researchers indicate that such

systems must be utilized naturally and must have been developed in line with certain procedures in the industry. However, there is no list of specific requirements regarding the type of database or prediction tool that should be used in their study. Useful suggestions are provided; nevertheless, they do not have anything to do with thyroid diseases or machine learning algorithms.

Biondi and Cappola (2021) focused entirely on the medical domain. This article examined the difficulty of interpreting hormone profiles, which is a complicated issue. There is no mention of machine learning, nor any discussion of metrics for predictions or automation. These results can be used directly by physicians, yet that is where their practical application ends.

One can cite the work of Esteva et al. (2021). They stressed the potential benefits of deep learning technologies in medicine, particularly in the area of image processing. They were able to achieve very high detection rates thanks to large databases and neural networks. However, this study did not consider laboratory data that could be structured, such as the information related to thyroids.

Sendak et al. (2021) studied the factors required to introduce machine learning to regular healthcare processes. They identified various issues in implementation such as trustworthiness, usability, and adaptability of models to particular infrastructure. This research is not based on particular dataset and predictive model but rather discusses general challenges faced by scientists trying to make machine learning usable in practice.

In 2021, Miotto et al. emphasized the significance of combining various data sets to enhance prediction power in the domain of medicine. The authors of the study surveyed various datasets and models and found that combination is the answer.

Moving away from the model-oriented research, Islam et al. (2022) examined different machine learning techniques for diagnosing thyroid diseases with the help of the UCI database. These researchers covered everything from data preprocessing to training a model and obtaining various measures such as accuracy and F1-score. The outcome was promising, yet the research lacked practicality in terms of physicians being able to implement it in their everyday practice. Similarly, Rajkomar et al. (2022) noted that while many models are technically sound, no one understands the logic behind them. In other words, the technology is available, yet not ready for implementation.

Jiang et al. (2022) considered the ways in which artificial intelligence might integrate into clinical procedures. These authors' approach relied mostly on theoretical considerations rather than practical solutions to be adopted by practitioners.

Back in Rajkomar et al. (2022), the authors revisited another important topic, i.e., besides being good at prediction, models should be interpretable and transparent. The topic was covered through a variety of medical applications, but thyroid disease and guidelines were not included. Nevertheless, the main idea was crystal clear.

Alshamrani et al. (2023) once again compared machine learning models that demonstrated high prediction rates based on the structured laboratory dataset related to thyroid diseases. However, similar to other works, the models remained rather a black box since they worked accurately but lacked interpretability. The same problem was identified by Shen et al. (2023).

Chiovato et al. (2023) brought back the basics of clinical work through their research in laboratories of the thyroid gland. Unlike the other works discussed above, Chiovato et al. (2023) focused on the problem clinically and noted that accurate diagnosis is quite difficult and complicated, especially without the use of AI technology.

Shen et al. (2023) carried out a detailed literature review on machine learning in predicting outcomes for endocrines. According to the findings of these researchers, it can be noticed that most researchers focus only on structured data and precision. There is no interest whatsoever in usability and explainability. This is the reason why more research is necessary.

Khan et al. (2024) showed once again that machine learning algorithms continue increasing the accuracy rate in diagnosing endocrines. As can be seen from other studies, researchers have not considered anything but structured data and traditional criteria.

2.5.1 Summary of Related

#	Study Name	Date	Author	Research Focus	Method / Approach	Key Contribution	Identified Limitation
1	American Thyroid Association Guidelines for Hyperthyroidism	2016	Ross et al.	Clinical diagnosis of thyroid disorders	Clinical guidelines based on laboratory biomarkers (TSH, T3, T4)	Established standardized diagnostic protocols for thyroid assessment	Does not include computational models or automated prediction
2	High-Performance Medicine: Human and AI	2019	Topol	AI in healthcare	Conceptual framework (no dataset or model)	Emphasized integration of human expertise with AI systems	No disease-specific implementation or predictive model
3	UCI Machine Learning Repository	2019	Dua & Graff	Dataset for ML research	Structured dataset for supervised learning (lab-based features)	Provides benchmark dataset for thyroid classification	Limited to laboratory data; lacks clinical context (Shen et al., 2023)
4	Clinical Decision Support Systems Review	2020	Sutton et al.	Decision-support systems	Literature review (no specific dataset/model)	Highlighted importance of interpretability and usability	No thyroid-specific application or ML implementation
5	Subclinical Thyroid Disease	2021	Biondi & Cappola	Thyroid physiology and clinical evaluation	Clinical review based on medical evidence	Provided comprehensive understanding of thyroid dysfunction	No computational or predictive modeling approach
6	Deep Learning in Healthcare	2021	Esteva et al.	AI in healthcare	Deep learning models (primarily imaging datasets) with accuracy-based evaluation	Demonstrated strong performance of AI in disease detection	Focuses on imaging; limited applicability to structured lab data; interpretability challenges (Miotto et al., 2021)
7	ML Challenges in Healthcare	2021	Sendak et al.	ML implementation in healthcare	Analytical review of real-world deployment challenges	Identified issues of usability, reliability, and integration	No disease-specific framework or predictive model
8	Deep Learning Review in Healthcare	2021	Miotto et al.	AI in healthcare	Literature review across multiple datasets and models	Emphasized importance of integrating multiple data sources	No structured framework for integration; interpretability remains limited
9	ML for Thyroid Disease Prediction	2022	Islam et al.	Thyroid classification	UCI dataset, ML models (DT, SVM, etc.),	Achieved strong classification	Focuses mainly on accuracy; limited

للإستخدام الداخلي - Restricted Internal use

					evaluation using Accuracy, Precision, Recall, F1-score	performance using lab data	interpretability (Rajkomar et al., 2022)
10	AI in Healthcare Review	2022	Jiang et al.	AI applications in healthcare	Review-based analysis (no dataset/model)	Highlighted importance of clinical integration of AI	No disease-specific implementation or technical framework
11	Machine Learning in Medicine	2022	Rajkomar et al.	ML in clinical practice	Review of ML applications in healthcare	Emphasized need for interpretability and clinical trust	Does not provide implementation methods for improving interpretability
12	ML Comparative Analysis for Thyroid	2023	Alshamrani et al.	Thyroid prediction	Structured dataset, multiple ML models, evaluated using classification metrics	Demonstrated effectiveness of ML algorithms for thyroid classification	Relies on lab-based features; lacks interpretability (Shen et al., 2023)
13	Advances in Thyroid Diagnosis	2023	Chiovato et al.	Clinical thyroid diagnosis	Clinical review of thyroid disorders	Highlighted importance of combining lab results with clinical understanding	No machine learning or predictive modeling approach
14	ML in Endocrine Diseases	2023	Shen et al.	ML applications in endocrine diseases	Systematic review of datasets, models, and evaluation methods	Identified trends in ML-based endocrine prediction	Most models rely on lab data and focus on accuracy; limited interpretability
15	AI in Endocrine Disease Prediction	2024	Khan et al.	ML in endocrine diseases	Comparative ML models with standard evaluation metrics	Showed potential of AI in improving prediction performance	Limited focus on interpretability and clinical usability (Sendak et al., 2021)

Chapter 3: Proposed Approach (Methodology)

Chapter 4: Experimental Results and Discussion

Chapter 5: Conclusion and Future Work

References

[Alshamrani, M., et al., 2023]

Alshamrani, Mohammed, Abdullah Alqahtani, and Fahad Alzahrani. “Machine learning approaches for thyroid disease prediction: A comparative analysis”. IEEE Access 11. 2023.

<https://scholar.google.com/scholar?q=machine+learning+approaches+for+thyroid+disease+prediction+comparative+analysis+2023>

[Biondi, B., & Cappola, A., 2021]

Biondi, Bernadette, and Anne R. Cappola. “Subclinical thyroid disease: Clinical relevance and management”. The Lancet Diabetes & Endocrinology 9. 2021.

<https://scholar.google.com/scholar?q=Subclinical+thyroid+disease+clinical+relevance+and+management+2021>

[Chiovato, L., et al., 2023]

Chiovato, Luca, Francesco Antonelli, and Paolo Vitti. “Thyroid disorders and their management: Recent advances in diagnosis and treatment”. Endocrine Reviews. 2023.

<https://scholar.google.com/scholar?q=thyroid+disorders+diagnosis+and+treatment+2023>

[Dua, D., & Graff, C., 2019]

Dua, Dheeru, and Casey Graff. "UCI Machine Learning Repository".

University of California, Irvine. 2019.

<http://archive.ics.uci.edu/ml>

[Esteva, A., et al., 2021]

Esteva, Andre, Alexandre Robicquet, Bharath Ramsundar, et al. "A Guide to Deep Learning in Healthcare". Nature Medicine 27. 2021.

<https://scholar.google.com/scholar?q=guide+to+deep+learning+in+health+care+2021>

[Islam, M. M., et al., 2022]

Islam, Md Mahmudul, Tanvir N. Poly, Buthaina Alsinglawi, Li-Fang Lin, Shu-Chen Chien, Jian-Cheng Liu, and Yu-Chuan Li. "Machine learning approaches for thyroid disease prediction: A comparative study".

Healthcare 10. 2022.

<https://www.mdpi.com/2227-9032/10/1/59>

[Jiang, F., et al., 2022]

Jiang, Fei, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Shanshan Ma, Yang Wang, Qiang Dong, Houbing Shen, and Yongjun Wang. "Artificial intelligence in healthcare: Past, present and future". Stroke and Vascular Neurology 7. 2022.

<https://scholar.google.com/scholar?q=Artificial+intelligence+in+healthcare+past+present+and+future+2022>

[Khan, M. A., et al., 2024]

Khan, Muhammad A., Saeed Anwar, and Fahad Alghamdi. “Recent Advances in Machine Learning Applications for Endocrine Disease Prediction”. Artificial Intelligence in Medicine. 2024.

<https://scholar.google.com/scholar?q=machine+learning+endocrine+disease+prediction+2024>

[Miotto, R., et al., 2021]

Miotto, Riccardo, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel Dudley. “Deep Learning for Healthcare: Review, Opportunities and Challenges”. Briefings in Bioinformatics 22. 2021.

<https://scholar.google.com/scholar?q=deep+learning+for+healthcare+review+opportunities+and+challenges>

[Rajkomar, A., et al., 2022]

Rajkomar, Alvin, Jeffrey Dean, and Isaac Kohane. “Machine learning in medicine”. New England Journal of Medicine 380. 2022.

<https://scholar.google.com/scholar?q=machine+learning+in+medicine+rajkomar>

[Ross, D. S., et al., 2016]

Ross, Douglas S., Henry B. Burch, David S. Cooper, Michael C. Greenlee, Peter Laurberg, Ana Luiza Maia, Scott A. Rivkees, Marvin Samuels, Julie A. Sosa, Marius N. Stan, and Martin A. Walter. "2016 American Thyroid Association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis". Thyroid 26. 2016.
<https://www.liebertpub.com/doi/full/10.1089/thy.2016.0229>

[Sendak, M., et al., 2021]

Sendak, Mark, William Gao, Nigel Brajer, and Suresh Balu. "Machine Learning in Health Care: A Critical Appraisal of Challenges and Opportunities". NPJ Digital Medicine 4. 2021.
<https://scholar.google.com/scholar?q=machine+learning+health+care+challenges+opportunities+sendak>

[Shen, J., et al., 2023]

Shen, Jian, Yuan Li, and Xiaolin Chen. "Applications of Machine Learning in Endocrine Disease Prediction: A Systematic Review". Frontiers in Endocrinology 14. 2023.
<https://scholar.google.com/scholar?q=machine+learning+endocrine+disease+prediction+review+2023>

[Sutton, R. T., et al., 2020]

Sutton, Russell T., David Pincock, Daniel C. Baumgart, David C. Sadowski, Richard N. Fedorak, and Karen I. Kroeker. “An overview of clinical decision support systems: Benefits, risks, and strategies for success”. NPJ Digital Medicine 3. 2020.

<https://www.nature.com/articles/s41746-020-0221-y>

[Topol, E., 2019]

Topol, Eric J. “High-Performance Medicine: The Convergence of Human and Artificial Intelligence”. Nature Medicine 25. 2019.

<https://scholar.google.com/scholar?q=high+performance+medicine+human+artificial+intelligence>