**Thyroid Disease Detection Using Multiple Machine Learning, Deep Learning and LIME Explainable AI( XAI) Techniques**

# Abstract

The Thyroid organ is a vascular organ what's more, one of the main organs of a human body. This organ secretes two chemicals which help in regulating the body's metabolism. The two Hyperthyroidism is one type of thyroid disorder. what's more, Hypothyroidism. At the point when this issue happens in the body, they discharge specific sort of chemicals into the body which uneven characters the body's digestion. Thyroid related Blood test is utilized to identify this illness yet it is frequently obscured and commotion will be present. Information purifying techniques were utilized to make the information crude enough for the examination to show the gamble of patients getting this infection. Various models, including Decision Tree, Random Forest, MLP, SVM, Logistic Regression, and LightGBM, alongside two deep learning models, ANN and 1D CNN, are utilized in our research project. We utilize standard experiment evaluations like precision, accuracy, recall, and F1-score to survey model execution, increased by Explainable AI (XAI) methods like LIME for interpretability. The outcomes show the predominance of ensemble techniques like Random Forest and XGBoost, close by interpretable models, such as Decision Trees, in precisely catching basic examples inside the dataset and making exact expectations. In particular, the Random Forest model accomplished an exactness of 0.92, accuracy of 0.72, review of 0.90, and an F1-score of 0.80. Likewise, the XGBoost model succeeded with an exactness of 0.92, accuracy of 0.84, review of 0.92, and a F1-score of 0.82. Besides, our investigation into profound learning procedures uncovers promising roads for working on indicative exactness, especially with ANNs, which exhibited a great precision pace of 0.96 and high accuracy and review values. Furthermore, the utilization of XAI procedures improves our understanding of model forecasts and element significance in thyroid sickness analysis. By tending to these roads, we expect to propel the field of thyroid sickness forecast, eventually adding to worked on persistent results and clinical decision-production processes.

***Keywords:*** Thyroid disease, machine learning, deep learning, predictive modeling, ensemble methods, Explainable AI, LIME, diagnostic accuracy.

**Table Of Contents**

[Board of Examiners I](#_Toc136692827)

[Declaration II](#_Toc136692828)

[Dedication III](#_Toc136692829)

[Acknowledgment IV](#_Toc136692830)

[Abstract V](#_Toc136692831)

[List of Figures IX](#_Toc136692832)

[List of Tables X](#_Toc136692833)

[List of Abbreviations XI](#_Toc136692834)

[Chapter 1: Introduction 1](#_Toc136692835)-5

[1.1 Background Of The Research 1](#_Toc136692837)

[1.2 Project Rationale 4](#_Toc136692838)

[1.3 Objective Of The Research 4](#_Toc136692839)

[1.4 Chapter Organization 5](#_Toc136692840)

[Chapter 2: Literature Review 6](#_Toc136692841)-9

[2.1 Research Area 6](#_Toc136692843)

[2.2 Major Challenges In Phishing Detection 6](#_Toc136692844)

[2.3 Studies Based On Deep Learning On Phishing Detection 7](#_Toc136692845)

[Chapter 3: Methodology 10](#_Toc136692846)-28

[3.1 Methodology 10](#_Toc136692848)

[3.2 Dataset Collection 12](#_Toc136692849)

[3.3 Data Visualization 13](#_Toc136692850)

[3.4 Correlation Heatmap 14](#_Toc136692851)

[3.5 Data Preprocessing 15](#_Toc136692852)

[3.5.1 Feature Scaling Using Normalization/Min-Max scaling 15](#_Toc136692853)

[3.6 Feature Selection Using Wrapper Method 16](#_Toc136692854)

[3.6.1 Forward Feature Selection 16](#_Toc136692855)

[3.6.2 Backward Feature Selection 16](#_Toc136692856)

[3.6.3 Exhaustive Feature Selection 16](#_Toc136692857)

[3.7 Train-Test Split 17](#_Toc136692858)

[3.8 Stratified K Fold Cross-Validation 17](#_Toc136692859)

[3.9 Training Of The Predicting Module 17](#_Toc136692860)

[3.10 Testing Of The Predicting Module 18](#_Toc136692861)

[3.11 Parameter Selection 18](#_Toc136692862)

[3.12 Deep Learning Models 19](#_Toc136692863)

[3.12.1 Multi-Layer Perceptron (MLP) 20](#_Toc136692864)

[3.12.2 Convolutional Neural Network (CNN) 22](#_Toc136692865)

[3.12.3 Long Short-Term Memory (LSTM) 23](#_Toc136692866)

[3.12.4 Hybrid CNN-LSTM 24](#_Toc136692867)

[3.13 Error Measurement 25](#_Toc136692868)

[3.14 Used Tools For The Research 26](#_Toc136692869)

[Chapter 4: Experimental Result Analysis 29](#_Toc136692870)-37

[4.1 Performance Evaluation And Results 29](#_Toc136692872)

[4.2 Train And Validation Loss 29](#_Toc136692873)

[4.2.1 Mean Squared Error (MSE) Loss 29](#_Toc136692874)

[4.3 Classification Report 31](#_Toc136692875)

[4.3.1 Precision 31](#_Toc136692876)

[4.3.2 Recall 31](#_Toc136692877)

[4.3.3 F1 Score 31](#_Toc136692878)

[4.3.4 Support 31](#_Toc136692879)

[4.4 Confusion Matrix 34](#_Toc136692880)

[4.5 Receiver Operating Curve (ROC) 36](#_Toc136692881)

[4.6 Result Analysis 37](#_Toc136692882)

[Chapter 5: Conclusion & Future Works 38](#_Toc136692883)-39

[5.1 Conclusion 38](#_Toc136692885)

[5.2 Future Works 39](#_Toc136692886)

[References 40](#_Toc136692887)-42

# List of Figures

**Fig. 3.1 Proposed System Architecture 12**

**Fig. 3.2 Percentage of Phishing And Legitimate Websites 13**

**Fig. 3.3 Histogram Plot Visualization 14**

**Fig. 3.4 Correlation Heatmap 15**

**Fig. 3.5 Multi-Layer Perceptron (MLP) Model 21**

**Fig. 3.6 Convolutional Neural Network (CNN) Model… 22**

**Fig. 3.7 Long Short-Term Memory (LSTM) Model 23**

**Fig. 3.8 Hybrid CNN-LSTM Model 25**

**Fig. 4.1 Training and Validation Loss 30**

**Fig. 4.2 Confusion Matrix of MLP 34**

**Fig. 4.3 Confusion Matrix of CNN 34**

**Fig. 4.4 Confusion Matrix of LSTM 35**

**Fig. 4.5 Confusion Matrix of CNN-LSTM 35**

**Fig. 4.5 ROC Curve 36**

**Fig. 4.7 Accuracy Comparison Of Proposed Deep Learning Models 37**

# List of Tables

**Table 3.1: The Distribution of Training and Testing Datasets 17**

**Table 3.2: Parameters 19**

**Table 4.1: Classification Report of MLP 32**

**Table 4.2: Classification Report of CNN 32**

**Table 4.3: Classification Report of LSTM 33**

**Table 4.4: Classification Report of CNN-LSTM 33**

**Table 4.5: Accuracy Comparison 37**

# List of Abbreviations

|  |  |
| --- | --- |
| **DL** | **Deep Learning** |
| **URL** | **Uniform Resource Locators** |
| **Machine Learning** | **ML** |
| **Multi-Layer Perceptron** | **MLP** |
| **CNN** | **Convolutional Neural Network** |
| **LSTM** | **Long Short-Term Memory** |
| **Mean Squared Error** | **MSE** |
| **TP** | **True Positive** |
| **TN** | **True Negative** |
| **FP** | **False Positive** |
| **FN** | **False Negative** |

# Chapter 1

# Introduction

## 1.1 Background of The Research

The reconciliation of computational science into the medical care industry has changed infection expectation and the executives by empowering the examination of huge patient datasets. Through the use of intelligent prediction algorithms, this development has greatly facilitated early disease detection. In spite of the wealth of clinical data frameworks and datasets, there exists a hole in the improvement of savvy frameworks prepared to do effectively dissecting sickness designs. AI calculations have arisen as urgent instruments in tending to this test, assuming an essential part in creating prescient models for different illnesses, including thyroid problems [1]. The thyroid organ, arranged in the neck beneath the Throat cartilage, assumes a crucial part in directing digestion and protein blend through the discharge of thyroid chemicals. These chemicals, to be specific levothyroxine (T4) and triiodothyronine (T3), apply command over indispensable physical processes, for example, pulse and calorie utilization. Iodine fills in as a basic part in the combination of these chemicals, with lacks or overabundances bringing about conditions like hypothyroidism and hyperthyroidism, separately [2][3].

**1.2 Project Rationale**

Thyroid sicknesses, enveloping circumstances like hyperthyroidism and hypothyroidism, have seen a flood in frequency rates lately. Given the thyroid organ's critical job in digestion guideline, irregularities inside it can appear in different anomalies, featuring the desperation of viable illness expectation and the executives systems. The finding and proactive therapy of thyroid diseases are basic to reduce expected complexities and work on tireless outcomes while restricting medical care costs. The IT technology progress in the area of data processing and estimating algorithms such as data mining and big data evaluation, as well the AI equivalence handling has open the way for a large numbers of eclectic applications in healthcare. Highly efficient methods and artificial intelligence are closely indispensable to modifying displaying a revolutionary outlook to fatal diseases prognosis. As diagnostic technologies are efficiently applied, healthcare practitioners can determine illnesses in the early stages, thereby enabling the made of personalized treatment plans and the specific use of medications befitting to the needs of each patient.

There are probably some problems behind despite the stepped-up progress in the illustration of illness by the artificial and deep learning models. Today's research is often guided by binary class problems, as people ignore the multi-class analysis and the normalized definition of components that are always part of the prediction of thyroid disease. Besides, the absence of execution examination between AI and profound learning models represents a constraint in improving prescient precision and model interpretability.

This study intends to address these holes by proposing an incorporated methodology for Thyroid Sickness Recognition using Different AI, Profound Learning, and LIME Reasonable man-made intelligence (XAI) Procedures. By utilizing these systems, the examination tries to foster a precise and interpretable prescient model for thyroid sickness finding, consequently working with proactive patient administration and further developing medical services results.

**1.3 Objectives of the Research**

The essential target of this examination is to foster a creative methodology for Thyroid Illness Identification utilizing a collaboration of Multiple AI, Deep Learning, and LIME Reasonable man-made intelligence (XAI) Procedures. The particular objectives of this study include:

* In the forth section, comparative analysis of multiple machine learning and deep learning models, ranging from Decision Trees, Logistic Regression, Support Vector Machines, Random Forests, LightGBM, Multi-Layer Perceptrons (MLP), Artificial Neural Networks (ANN) and Deep Neural Networks (DNN) is conducted.
* Aggregating and preprocessing a diverse dataset encompassing demographic and medical features associated with hyperthyroidism, drawing from reputable sources such as Kaggle and medical repositories.
* Evaluating the performance of the developed models using standard evaluation metrics, including accuracy, precision, recall, and F1 score, to ascertain their efficacy in thyroid disease prediction.
* Implementing the LIME approach to provide interpretable explanations for the predictions generated by the deep learning model, thereby enhancing transparency and trust in the diagnostic process [3].

By achieving these goals, the examination attempts to spearhead the improvement of an exact, versatile, and interpretable prescient model for thyroid infection conclusion. Such headways hold significant ramifications for medical services conveyance, engaging clinicians with opportune experiences and empowering customized therapy methodologies custom-made to individual patient necessities.

**1.4 Chapter Organization**

This research is structured into five cohesive sections:

* Chapter 1 serves as the introductory chapter, providing an overview of the research subject, rationale, and objectives.
* Chapter 2 delves into a comprehensive review of existing literature pertinent to thyroid disease diagnosis, machine learning techniques, and the integration of artificial intelligence in healthcare.
* Chapter 3 elucidates the methodological framework employed in this study, encompassing data collection, preprocessing techniques, model development, and evaluation methodologies.
* Chapter 4 presents the empirical findings and analysis derived from the implementation of the proposed approach, highlighting key insights and performance metrics.
* Chapter 5 concludes the research endeavor, synthesizing the findings, delineating implications for healthcare practice, and outlining avenues for future research and innovation.

# Chapter 2

# Literature Review

**2.1 Research Area**

The thesaurus words "Thyroid illness", "Thyroid cancer", "AI", and "deep learning" were used as an avenue to select the essential articles. On the contrary, the amount of salvaged results becomes completely other practice for identification the most important ones. Thus, we have increased the request questions inquiries and used very close sorting. First, let us say that more than 100 vital articles were found by the talk itself. The first task we performed was a literature review on those articles and then selected 25 of them which completely relate to our research. In this approach, AI (Artificial Intelligence) and deep learning strategies are applied for thyroid disease ID and thyroid malignancy detection. In this instance the in-depth procedure of a case is being compared with the one that is viewed at the higher scale. Thus the method that is used in two co-related activities is examined distinctively. [1]Ankita Tyagi, Rikitha Mehra "Intelligent Thyroid Infection Expectation Framework Utilizing AI Techniques" at fifth InternationalEEE Global Meet on Equal, Shared and Distributed Computing(PDGC-2018), 20th-22nd of December, 2018, Solan, India. The technique they used for the different types of calculation include Order Decision Tree, Support Vector Machine, Artificial Neural Organization, k-nearest Neighbor calculation. What the study will do is inputting UCI Vault information, clustering it, expecting the result, and then getting the accuracy by means of estimation. They have used this money to begin to examine and define precision of equations and there has been a search for the most accurate method.Sunila Godara. [3] They have employed Logistics Regression and SVM AI Methods in the crunch of the Thyroid Data. Correlation was calculated between the precision, recall & F measure, ROC (accuracy) & RMS (root mean square) errors. As it turned out, Logistic Regression showed excellent classification performance.YongFeng Wang. [2] Ultra sound images of the thyroid organ are reviewed either as harmless or threatening type in order to under go radiomics and deep learning approaches. Exploring the parallels of these approaches is the focus. After applying both strategies, the results in terms of characterization accuracy, awareness, and particularity gets 66.81%, 51.19% and 75.77%, correspondingly, which are then followed up by 74.69%, 63.10% and 80.20%, accordingly, with the testing files for the deep learning based technique. This types of approaches were proved to be the best ones. Hitesh Garg. [4] Feed Forward Neural Architecture is applied to the Ultrasound images for the feature extraction and segmentation of Cancers. While, accuracy and different elements were similar and every factor was high of 86%. The authors of the concentrate by Li et al. [6] were the one who implemented the Rope (Linear shrinkage model) and LR model in selecting the variants of the observational features of the compromising thyroid level that was related with ultrasound image. Subsequently, RF is applied proximate to a scoring system which assesses (thyroid goiter) as a dangerous handle. Among the methods, the logistic rope regression (LLR) with RF got the highest result with a hit rate of 82% Kim and colleagues conducted AI-based forecast of BRAF mutation presence in thyroid mutated human handles as well (Ref. 7). To carry out this study, the creators selected 96 final images of true thyroid gland via ultrasonic examination. Among all the images, 86 features were solved by the algorithms, and the LR, SVM, and RF models were used in predicting whether BRAF change exists or not. What we can read from the data is that each model accomplished top accuracy that stood at 64.3%. One of the clinic tools that utilize AI is animitimize false negative diagnosis in the first stages of thyroid cancer. It is done by fine-needle goal (FNA) element and ultrasonic aspects because of AI-based thyroid knob harm prevention [8]. RF secure better performance when compared to various other techniques such as decision tree (DT) and slope down (GD). To some incompleteness of the presented works' appearance, the interpretation of the thyroid disease discovery isn't totally clear despite the probable space for the further improvement.

**2.2 Major Challenges in Thyroid Disease Detection**

A few thyroid disease identification and classification approaches have been introduced in the writing. For instance, Garcia et al. [9] anticipated the high conceivable iotas beginning the thyroid chemical homeostasis using AI estimations RF, LR, GBM, SVM, and deep neural networks (DNN). The early assumption for the particles is valuable for extra testing in the essential periods of thyroid disease. The sub-nuclear events were gotten from ToxCast datasets for running the trials. The article detailed that Thyroid Peroxidase (TPO) and Thyroid Chemical receptor (TR) achieved the best insightful execution with a F1 score of 0.83 and 0.81, independently. The makers in [10] utilized the image handling methods and component choice strategies to pick the significant features from the dataset and achieve the best show for thyroid sickness assumption.

**2.3 Studies Based on Machine Learning On Thyroid Disease Detection**

Razia's et al [11] narrow down that it is either excessive or less hyperthyroidism. Besides, they broaden up that it is also hypothyroidism. There are four data sets that were the outputs of UC Irvine data collection system and were made public. The end is made up of 7200 cases with the forecasting probability for the existence of cardiovascular disease. Provide necessary information for each case concerning the 21 population criteria mentioned above. According to the author this algorithm was better than the other two algorithms by a huge margin as it had a prediction power of 99.23%. In most cases, those multi-functional layers only contains a few types of data modes and very limited information as preprocess, thus their solutions won't work with all continuous data. The disease of the thyroid origin, which is known as the grouping problem, is going to be answered with multi-bit SVM using [12]. Developers reported that 97.49% accuracy on average was achieved on the thyroid datasets as major parameters in the multi-stage SVM processing were assessed. As an illustration of a wolf in the environment, whereby, there is multiple features selection and improvement, which also gambit better service quality, will be observed as a result of a better mode of transportation.

**2.4 Studies Based On Deep Learning On Thyroid Disease Detection**

A survey [13] performed multiclass hypothyroidism using explicit highlights and AI computations. Hypothyroidism is requested into four classes. The results show that RF performed well with 99.81% accuracy contrasted with the SVM, KNN, and DT computations. Nonetheless, the makers didn't make reference to the introduction of their proposed methodology for thyroid contamination arrangement. Another survey [14] tried three component selection strategies close by SVM, DT, RF, LR, and Guileless Bayes (NB) to make early assumptions for hypothyroidism. Three element selection strategies, recursive element selection (RFE), univariate highlight selection (UFS), and head part examination (PCA), are tried in blend in with ML estimations. The RFE mix with ML estimations performed better contrasted with other element selection strategies. All the five ML computations gained 99.35% accuracy when gotten together with RFE incorporate selection. Nonetheless, the information test size is small, with only 519 records. A colossal extension dataset is supposed to survey the practicality of their strategy. The makers [15] evaluated the show of the thyroid disorder portrayal using different AI estimations. SVM, RF, DT, NB, LR, K closest neighbor (KNN), and MLP are used for infection assumption. A dataset trial of 1250 is taken from crisis centers and exploration offices in Iraq. The MLP anticipated the thyroid portrayal with 96.4% accuracy. Notwithstanding, there is still space for execution improvement. Hosseinzadeh et al. [16] proposed a multiple multi-layer insight (MMLP) strategy to portray thyroid diseases. Exactly when the MMLP is applied close by a lot of six networks, the accuracy is chipped away at by 0.7% contrasted with a lone MLP. In spite of the way that MMLP procured near 100 percent request accuracy on gigantic dataset tests, planning deep learning strategies like MMLP is costly and needs high computational resources for get ready speedier. The KNN with various distance abilities is executed to test the thyroid ailment recognizable proof in [17]. The chi-square and L1-based highlighted selection procedures were used to pick the ideal elements preceding applying the KNN with Euclidean and Cosine distances. The makers detailed that KNN obtained promising results. Anyway, the tried model size is small, with 590 models out and out.

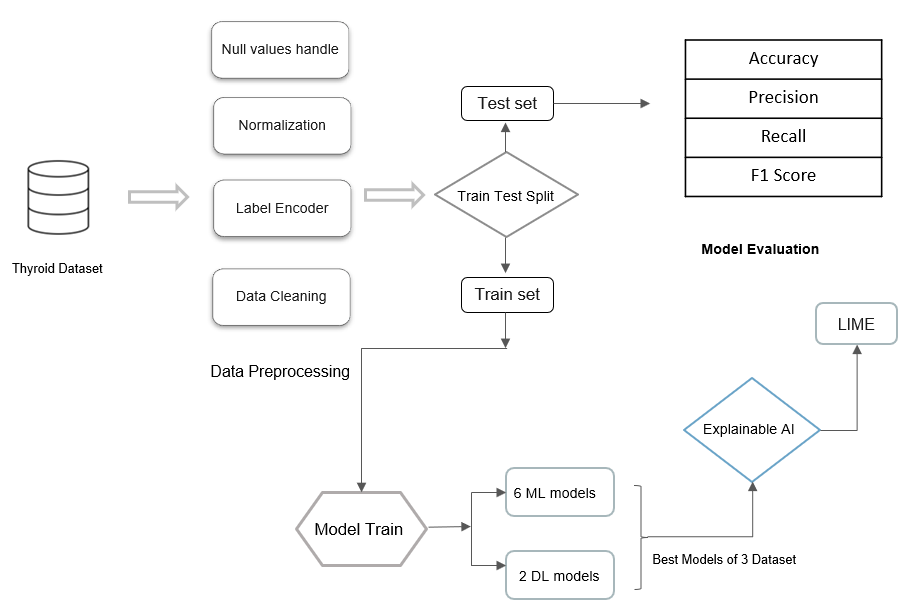
Mishra et al. [18] applied the ML strategies successive insignificant streamlining, Decision tree, Random Forest, and K-star classifier to foresee hypothyroid infection. A model size of novel 3772 records is considered for this survey. The makers announced that RF and DT performed better contrasted with the following two techniques, with accuracy scores of 99.44% and 98.97%. Nevertheless, the makers didn't consider hyperthyroid assumption. Alyas et al. [19] played out a close to examination of the AI strategies DT, RF, KNN, and artificial neural organization (ANN) to recognize thyroid infection. The tests were driven on the greatest dataset and considered both inspected and unsampled information for thyroid affliction assumption. RF procured the best assumption with 94.8% accuracy. Nonetheless, the makers didn't play out the thyroid disease type assumption tests. Experts furthermore applied deep learning models to anticipate thyroid ailment portrayal. For instance, the makers [20] used a deep neural organization (DNN) to foresee the thyroid infection portrayal. The show evaluation is done on the UCI dataset of 3152 noteworthy models. The makers detailed 99.95% accuracy while using DNN to describe thyroid disease. Nonetheless, a tremendous dataset is supposed to set up the model for execution evaluation properly. Likewise, extra figuring resources are supposed to set up the deep learning models.

In the last years, there has been a ton of work done to analyze the discrete illnesses in thyroid. Many creators have utilized different sorts of information mining procedures. The creators demonstrated to acquire a sufficient methodology and conviction to figure out the illnesses comparable to the thyroid by the work that incorporates different datasets and calculations connected with the work that will be finished later on viewpoint to achieve powerful and improved results. The aim of the paper deciphers different strategies of information mining components and the factual qualities that are been promoted in the last a very long time for understanding of thyroid illnesses with sureness by different creators to accomplish different possibilities and for different methodologies. There are different calculations of AI counting random woods, decision tree, credulous Bayes, SVM and ANN that are broadly utilized in the regular illnesses and in the prognostic issues. There are not many capabilities that are involved infections connected with coronary illness, diabetes, Parkinson's, hypertension, the Ebola virus(EV), findings, and gauging, R-NA sequenced information investigation and designation of biomedical imaging. Regardless of the progression of an AI set illness expectation system and a clinical assurance is a nontrivial task. There is fundamental issues for example securing of information, gathering and gathering that are worn to prepare the AI structures. In the real movement issues, assessment of huge informational collections in biomedical over a deep continuation are wanted, and are basically non-existent. In a deliberate methodology for prior conclusion of Thyroid illness utilizing back proliferation calculation utilized in neural organization. ANN fragile and lays out on back proliferation of a blunder that is being utilized for earlier illness expectations. The effect of ANN is being prepared with the exact information and testing systems that are borne out as information that were not being used during the most common way of preparing. ANN finishes up in great consistence with the starter information and demonstrates the high level neural organization which utilizes as a substitute for earlier illness expectations. In the creators examined and look at the four grouping models specifically Gullible Bayes, Decision Tree, Multilayer Perceptron, and Outspread Premise Capability Organization. The determination exhibits a pivotal accuracy for all the characterization models. The Decision Tree model surpasses by the other grouping models. In this work, 29 credits of the dataset are recruited and implemented as an Element Selection procedure for example Chi-Square, The datasets are being separated by directing the solo covered channels on the traits for change in the nonstop qualities into ostensible and thus lessen the 29 ascribes to 10 credits. AI (ML) is a division of artificial knowledge and is penetrated in the components of logical exploration at developing advances. AI works with calculations to audit for a fact without quite being focused on. AI has been actuated by the information explosion that is associated with an expanding computational ability, and traditional the study of disease transmission are a high level mixed late information science way to deal with lash the capacities of the refined information. To consider immense game plans of information, the specific apparatuses investigate in neighboring clinically pertinent contact among information and result model. Verifiable investigations of careful ends are famously deceivable to alter careful agreements. Conclusive parts of careful agreements are depiction of the patient's confidant that guides from a medical procedure in the mediation. AI empowers PCs to decide from going before information to make careful expectations on current information. The enlightening aspect makes exceptionally legitimate forecast calculations that can duplicate the previously extraordinary correspondence in immense, tangled sets of information and adapt to viable information emanation. The composite qualities and the remedial methodology that are being utilized in the thyroid issues cater a more than adequate grouping of many-sided and grouped information and subsequently, a favorable structure for the plan of AI models. This proposes an adequate plausible for the usage of AI models and supports a thriving propensity towards thorough medications in which therapeutics are sewn to specific patients. In the field of AI, a broad disparity could be created in the midst of regulated and unaided learning. Directed gaining calculations decide from "marked" preparing information to edit a model that achieves expectations on previously fanciful information. For unaided system of learning, just unlabeled information are attainable and the calculations looks to resource the relationships and gadgets, solo learning calculations might get the huge number of unlabeled genomics information as information and examine previously unknown array of information. These calculations may, in some way or another, be prevailing in already previous plans in complex information that are not principally quantifiable by people and might be utilized to foster names to prepare a managed model at long last. In regular programming, a software engineer physically makes a bunch of data - "the projects" - to create a want yield from a given arrangement of information factors. In AI, the sources of info are outfitted along with the ache for result and PC calculations are asked to determine the "rules from the arranged preparation information". An electronic educational experience is a satisfactory approach to deciphering a tremendous overflow of information, planning covered correspondences in composite arrangements of information, and alluring to dynamic emanation. In the learning system, calculations try to resource the magnificent collection of info factors (elements), and loads are incorporated into these highlights in the model by reducing the uniqueness between the expected and significant outcomes. AI is utilized in preparing the framework over huge data sets, where the upheld AI procedures are reused to foster reflection gadgets or edge a model and utilize the refined gadgets or casing a model and utilize the cultivated gadgets or models in making expectations coming down the line for mysterious cases.

# Chapter 3

# Methodology

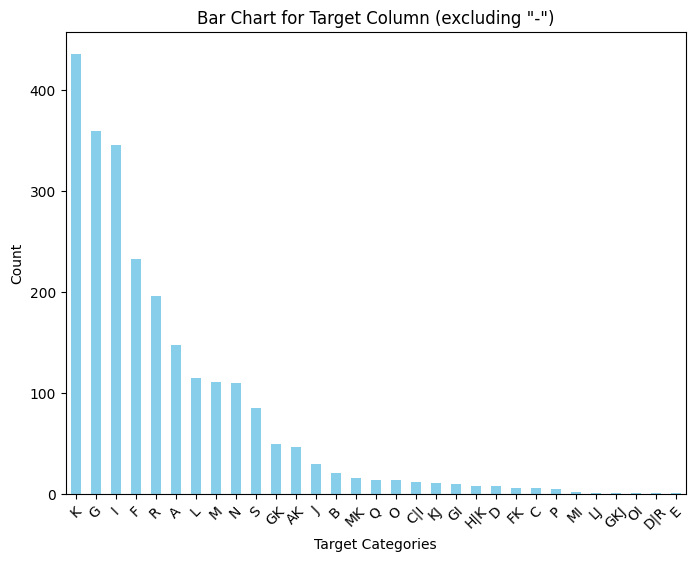
Different AI techniques are evaluated during model preparation, and the best one is picked. Thus, the framework's effectiveness will get to the next level. The properties chose depend on wellbeing aptitude. Cleaned information is used to prepare and test the calculation. The calculation gathers qualities from multiple datasets and characterizes them in view of names. Test information is given into the framework to guarantee conjecture accuracy. The gained highlights will be contrasted with give a likelihood for the test information. The most elevated likelihood worth will be doled out to a specific name, regardless of whether Thyroid. Figure 1 shows the by and large calculated plan of the framework that is being proposed. The parts of the proposed framework are depicted as follows.



**Figure 01:** Overall Research Methodology.

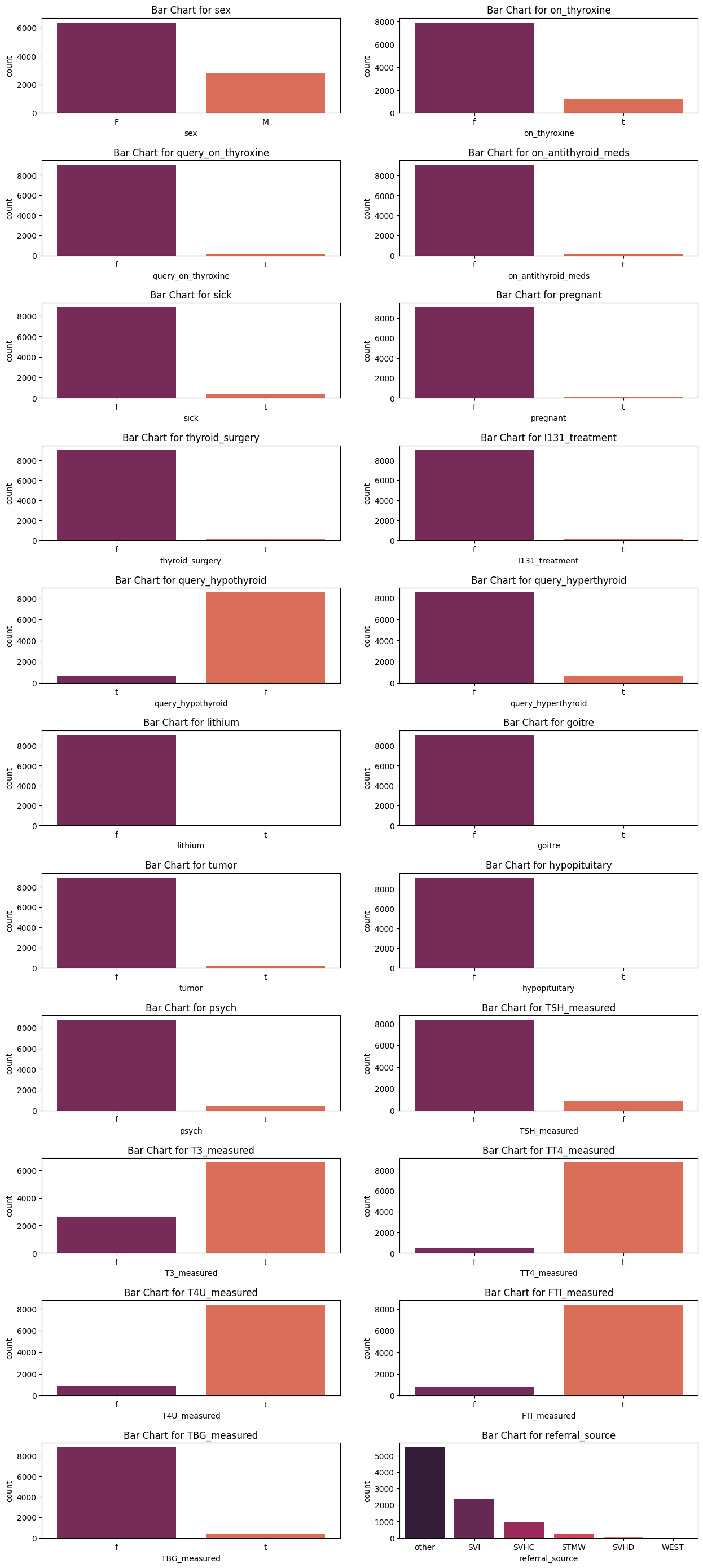
**3.1 Dataset Details:**  The dataset is collected from Kaggle where there are total of 9172 rows and 31 features. There are also plenty of missing values which should be handled. This dataset is created by Ross Quinlan from the Garavan Institute of Australia. Our dataset features name with their description are given below:

|  |  |
| --- | --- |
| **Feature Names** | **Value Type** |
| age (int) | continuous |
| sex (str) | Male, Female |
| on\_thyroxine (bool) | False, True |
| query on thyroxine (bool) | False, True |
| on antithyroid meds (bool) | False, True |
| sick (bool) | False, True |
| pregnant (bool) | False, True |
| thyroid\_surgery (bool) | False, True |
| I131\_treatment (bool) | False, True |
| query\_hypothyroid (bool) | False, True |
| query\_hyperthyroid (bool) | False, True |
| lithium (bool) | False, True |
| goitre (bool) | False, True |
| tumor (bool) | False, True |
| hypopituitary (float) | False, True |
| psych (bool) | False, True |
| TSH\_measured (bool) | False, True |
| TSH (float) | continuous |
| T3\_measured (bool) | False, True |
| T3 (float) | continuous |
| TT4\_measured (bool) | False, True |
| TT4 (float) | continuous |
| T4U\_measured (bool) | False, True |
| T4U (float) | continuous |
| FTI\_measured (bool) | False, True |
| FTI (float) | continuous |
| TBG\_measured (bool) | False, True |
| TBG (float) | continuous |
| referral\_source (str) | WEST, STMW, SVHC, SVI, SVHD, other |
| target (str) | continuous |

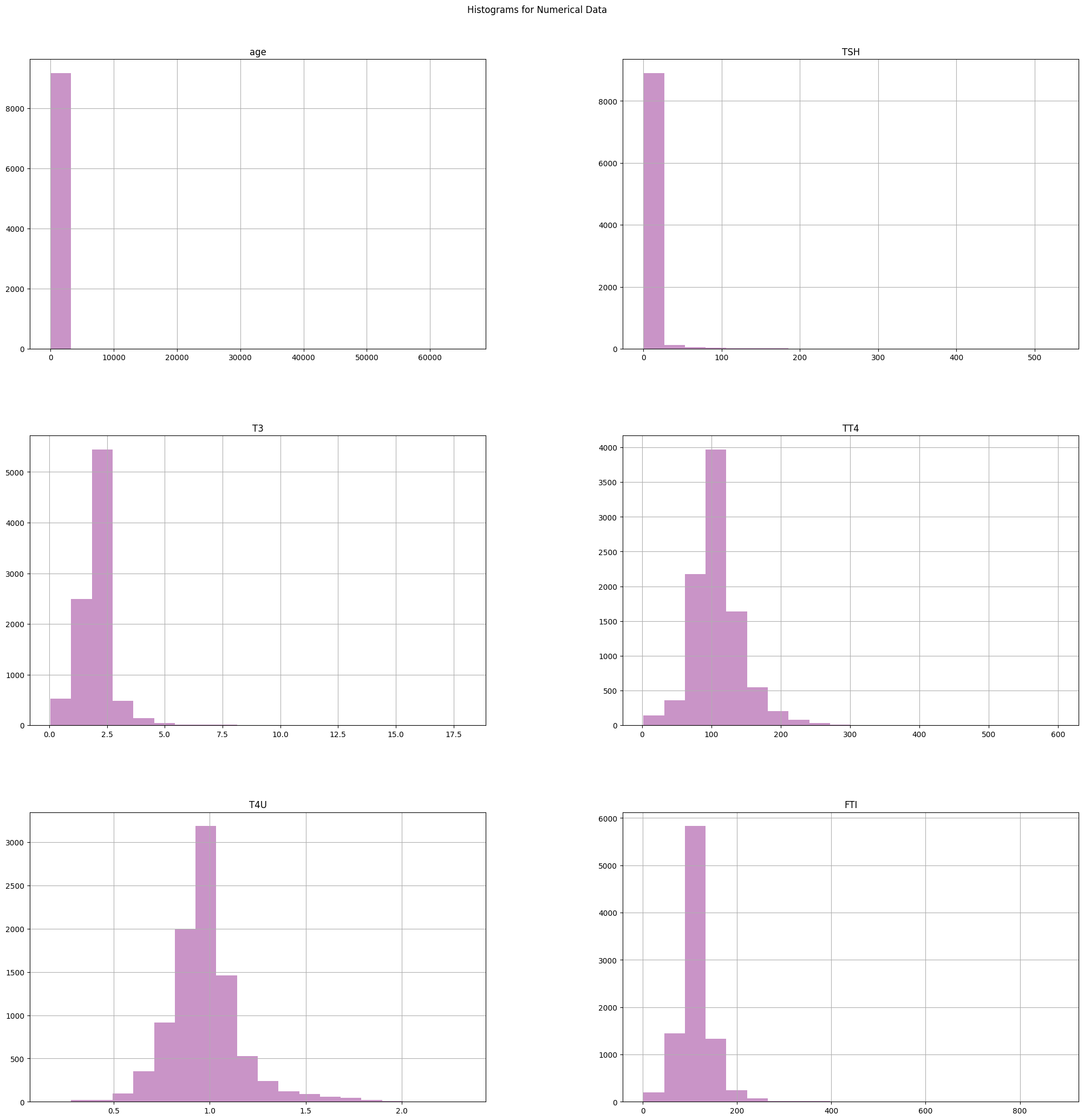
****

**Figure: Target class distribution.**

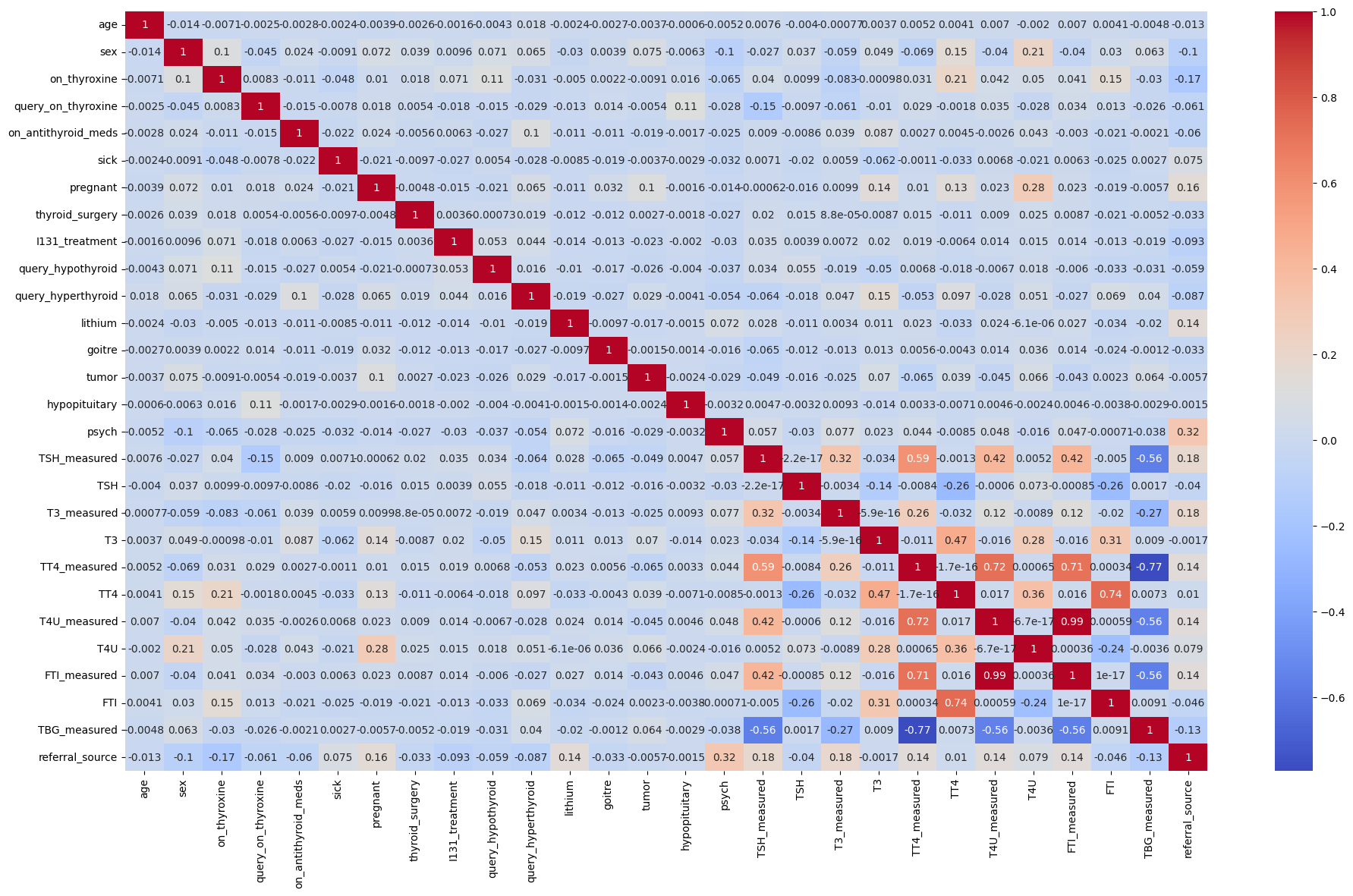
**3.2 Data Pre-processing:** Preprocessing data is necessary to ensure it is intelligible. Before preprocessing the data, check for missing values. Replace missing values with the feature's mean, medium, or mode. The category data has to be transformed into numerical data. To use machine learning algorithms, the dataset is divided into training and testing sets. Correcting missing numbers and removing unnecessary information can enhance overall accuracy. Missing values must be processed to avoid losing critical information and badly impacting outcomes. Next, feature scaling using the min-max approach is used to get the highest and lowest entry values. Developing a classification model requires a high-quality train dataset for accurately predicting potential outcomes. The four main tasks in Classification Predictive Modeling include Binary Classification, Multiple-Class Classification, Multi-label classification and Imbalanced classification. A binary classification consists of only two class labels. Some examples are Decision Tree, Logistic Regression, and Naïve Bayes. Multi-Class Classification includes more than two class tags. This classification method is effective with a large number of classifiers. Binary Classification algorithms can also be used for multi-class classification. Random Forest is a good illustration of this sort of classification assignment. Multi-Labeled Classification utilizes several class tags, with one or more able to predict the outcomes of another. Examples include multi-label decision trees and multi-label random forests. Imbalanced Classification refers to classification sets where the number of example classes is not evenly distributed. Our dataset’s target feature is multi-class classification based which consists of 10 different class type.



**Figure**: Bar chart for categorical features.



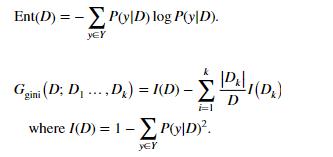
**Figure**: Histogram for Numerical Features.



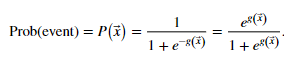
**Figure**: Correlation Matrix.

**3.3 Machine Learning Models:** After train test split, the dataset is divided into 2 parts – Train set (80%) and Test set (20%). For transforming all the numerical features into same ranges of 0 and 1, we used standard scaling technique. Also, we used Label Encoder for transforming all categorial features into numerical one. Then we have applied 6 different machine learning models and 2 deep learning models in this dataset. 6 of the machine learning models are – Decision Tree, Random Forest, MLP, Support Vector Machine, Logistic Regression and LightGBM.

A decision tree algorithm uses a top-down strategy to develop its structure. The ID3 algorithm is used to create the decision tree. Eliminating unnecessary elements increases classification accuracy. The decision tree algorithm calculates values from the thyroid patient data. The computation is based on the training dataset. Increasing the number of records in the training dataset improves the algorithm's accuracy. Decision trees are a notable way to decide. The methodology of 'partition and overcome' includes building decision zones by parting the occasion space. Subsequent to testing, a root hub is framed. The dataset is then separated in view of the worth of the comparing test property. The class is addressed by a leaf hub toward the finish of a tree. The decision is not entirely set in stone by the hub's branch or course.



Logistic regression (LR) is a classification model in machine learning, ordinarily used in spaces like health and sociology. Logistic regression is utilized in different examinations to distinguish risk factors and conjecture ailment probabilities. These forecasts are discrete, demonstrating exact qualities or gatherings. Clients can analyze likelihood evaluations for the model's classifications.



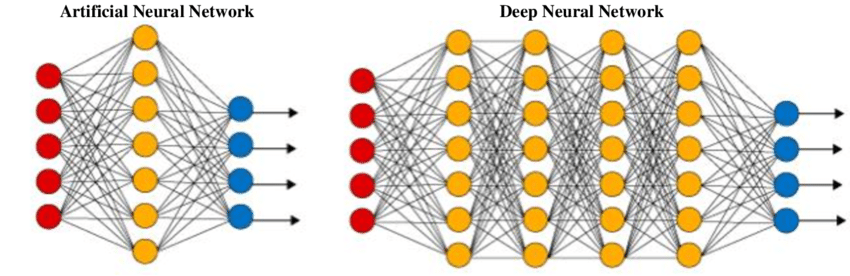
SVM is a managed machine learning strategy fit for leading classification, regression, and exception ID. The dataset's attributes are displayed in n-layered space. The two classes separate by defining a straight boundary known as a hyperplane. The SVM technique chooses a line that separates two classes while keeping away from the closest examples. The "support vector" in "support vector machine" alludes to two position vectors from the beginning to the decision limit focuses.

The Random Forest algorithm is a viable tree learning approach (Ensemble Learning) in Machine Learning. It works by producing various Decision Trees during the preparation stage. Each tree is worked by taking a random subset of the informational collection and estimating a random subset of qualities in every division. This capriciousness changes up individual trees, bringing down the risk of overfitting and boosting by and large expectation execution. In expectation, the algorithm consolidates the results, everything being equal, either by deciding in favor of (classification undertakings) or averaging (for regression assignments). This cooperative decision-production strategy, supported by many trees' bits of knowledge, exhibits consistent and careful outcomes.

A LightGBM tree is a decision tree structure used by the LightGBM gradient boosting framework. Hubs reflect highlight parts, while leaf hubs incorporate expectations. LightGBM trees are assembled recursively, leaf-by-leaf, with the objective of maximal misfortune decrease at each preparing step. Each split plans to expand a particular goal capability. It gives different dividing rules and pruning approaches for working on model execution. These trees contain a troupe model, which predicts by collecting the results of multiple trees, bringing about exact and proficient machine learning models.

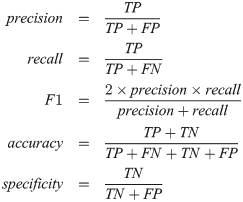
Classification utilizing Multi-Layer Perceptrons (MLP) is a significant machine learning approach in the counterfeit brain network class. It is a flexible and effective methodology for tending to an extensive variety of classification issues, including text classification and picture acknowledgment. Traditional linear classifiers may not be capable, yet MLPs are perceived for their capacity to address mind boggling, nonlinear associations in information. Here, we will take a gander at how to use the famous Python machine learning framework scikit-figure out how to build classification with MLPs.

ANN comprises an input layer, a couple of hidden layers, and an output layer. While these networks can handle basic numerical issues and PC issues, for example, essential entryway designs and their comparing truth tables, they battle to perform complex picture handling, PC vision, and normal language handling errands. For these difficulties, we utilize profound brain networks, which habitually highlight a muddled hidden layer structure with many layers, including a convolutional layer, max-pooling layer, dense layer, and other unmistakable layers. These additional layers permit the model to distinguish hardships and convey ideal answers for muddled undertakings more readily.



**Figure**: ANN and Deep Neural Network difference.

We have evaluated accuracy, precision, recall and F1 score for these models.



**Explainable AI:** We employed well-known explainable AI approach, LIME, to create local and global explanations for the deep learning model's validation and test predictions. LIME creates an interpretable model by training a local linear model around the prediction point. LIME is a well-known model independent approach for explaining individual predictions made by a black-box model. It constructs a local linear model around the prediction point and weights the input characteristics to determine their significance in the prediction. We utilized Python's Lime module to provide explanations for our model's predictions on the validation and testing sets. LIME delivers locally consistent explanations in the neighborhood of the instance being described. By default, it generates 5000 samples of the feature vector using the normal distribution. The prediction model, whose decisions are being explained, is then used to retrieve the target variable for these 5000 samples. After acquiring the surrogate dataset, it weights each row based on how similar it is to the original sample/observation. Then it employs a feature selection approach such as Lasso to retrieve the most significant characteristics. LIME additionally applies a Ridge Regression model to the samples, utilizing just the acquired features. The output forecast should be similar in magnitude to the one produced by the original prediction model. This is done to emphasize the relevance and significance of the discovered attributes.

# Chapter 4

# Experimental Result Analysis

**4.1 Performance Evaluation and Results**

In our review, we assessed thyroid sickness forecast utilizing six machine learning models: Decision Tree, Random Forest, MLP, SVM, Logistic Regression, and LightGBM. and two profound learning models: ANN and 1D CNN.We surveyed their performance utilizing standard assessment measurements like accuracy, precision, recall, and F1 score. Moreover, we utilized LIME for explainable AI, giving bits of knowledge into the models' expectations.

## 4.2 Machine Learning Models as a solution

In this segment, we lead a careful analysis of the viability of different machine learning models in foreseeing thyroid illness, assessing their performance through key measurements like accuracy, precision, recall, and F1-score. This analysis reveals insight into the models' abilities to explore the intricacies of the dataset actually. Among the variety of models analyzed, three stand out for their predictable greatness across multiple assessment boundaries. The Random Forest model is especially vital for its outstanding robustness, bragging an accuracy 0.92, precision of 0.72, recall of 0.90, and a F1-score of 0.80. Its ensemble method, which capably designs the perplexing data joins while definitely decreasing the gamble of overfitting, gives it its solidarity. Interestingly, the XGBoost model performs well with 0.92 accuracy, 0.84 precision, and 0.92 recall, which is very well. Its gradient boosting approach, which definitively changes its antipathy for the unobtrusive qualifiers inside the dataset, is credited with its prosperity. This outcomes in prevailing execution across undeniably estimated measurements. Besides, the Decision Tree model exhibits astounding ampleness and is notable for its straightforward interpretability. It accomplishes an accuracy of 0.93, though the precision is 70% and recall is 69%, and a F1-score is 69%, which shows express strength in accuracy and recall by and large. A nitty gritty examination of individual estimations reveals that these driving models beat others in achieving the most raised scores in accuracy, precision, and recall. They exhibit talented at summarizing the fundamental models, restricting false up-sides, and definitively perceiving a serious degree of certified positive cases. Alternately, models like the Support Vector Classifier (SVC), Logistic Regression, and LightGBM show somewhat more delicate execution across these estimations. This may be a direct result of their challenges in getting mind boggling models inside the dataset or their repugnance for racket and exemptions. Figure 3 addresses the accuracy of the different models, giving a visual comparison of their exhibition.

A graph showing different colored rectangular shapes

Description automatically generated

**Figure 3 : Accuracy of the different models**

The examination of the exhibition of various AI (ML) models in anticipating thyroid illness gives important experiences into the adequacy of various calculations and their reasonableness for this particular prescient errand. The discoveries feature a few central issues in regard to show execution, interpretability, and generalizability. Table 01 Shows the aftereffects of various assessment measurements of AI models.

**Table 01: Result analysis of Machine Learning models**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1 score** |
| Logistic Reg. | 75% | 65 | 75 | 66 |
| LightGBM | 61% | 63 | 61 | 62 |
| MLP | 84% | 82 | 84 | 82 |
| SVC | 75% | 61 | 75 | 65 |
| Random Forest | 92% | 92 | 92 | 92 |
| Decision Tree | 93% | 93 | 93 | 93 |

The Random Forest, XGBoost and Decision Tree models are seen to outperform the others in all the major statistics showing emergence of more hidden designs inside the dataset as compared to other models. These models are based in apparel and prove to be more exact, more specific and more accurate which shows their capability to slim down and make reasonable inferences. XGBoost and Random Forest model merge strikes the finest balance between preventing overfitting and not missing the various connections within the information space. On the contrary, its simplicity and originality are prime factors that help it achieve many significant elements and relationships, which again makes its performance very robust. My other hand, the SVC and LightGBM models compete for the accuracy as the other side, especially their precision and recall are not so good. The conventional SVC tries to find a valid anchor point inside the data which will eventually lead to sub-standard decision-making and the failure to catch examples. Besides, with regard to Large datasets handling, LightGBM is the productive one, its ability to tackle different cases, however, fails to give an explanation why that actually happens. With deeper analysis of leading models, their characteristics which constitute their strengths are clearly revealed. With the Random Forest model, the group approach helps to combine forecasts from various decision trees, hence increasing the accuracy of prediction as well as the strength to withstand abnormalities and complex data situations. Also, the gradient boosting method of the XGBoost model aims to successively revamp the weak models to target the errors, hence its high recall and accuracy rate is intelligible. Moreover, the Deterministic model's manageability and transparency facilitates the alignment of it to the most significant patterns underlying the data sets, thus, attaining high precision. Finally, the performance of the Random Forest, XGBoost, and Decision Tree models has been found highly valuable in detecting the risk of thyroid disease, providing the upper hand in all the evaluation indices. This just validates the significance of selecting the best machine learning algorithm which performs exemplarily for the type of data and the task at stake. Furthermore, data mining and model analysis can be carried out and design of features and components may be tried which gives extra information with better prediction accuracy, and finally helps to increase the effectiveness of forecast models of thyroid diseases and also clinical decision-making procedures. Predictions 4, 5, and 6 are the chaos networks for Random Forest, XGBoost, and Decision Tree models separately.

A sheet of a paper with numbers

Description automatically generated

Figure 4: Confusion matrix of Random Forest

A sheet of a math test

Description automatically generated with medium confidence

Figure 5: Confusion matrix of XGBoost

A sheet of a paper with numbers

Description automatically generated

Figure 6 Confusion matrix of Decision Tree

These outcomes underscore the significance of picking machine learning algorithms that are well fit to the particular qualities of the dataset and the expectation job needing to be done. In particular, in the domain of thyroid sickness expectation, it becomes apparent that gathering strategies, for example, Random Forest and XGBoost, close by additional interpretable models like Decision Trees, stand out as far as accomplishing prevalent prescient performance. Moreover, the capacity to decipher these models holds specific importance in medical settings. This is on the grounds that appreciating the rationale behind the expectations is basic for pursuing informed clinical choices.

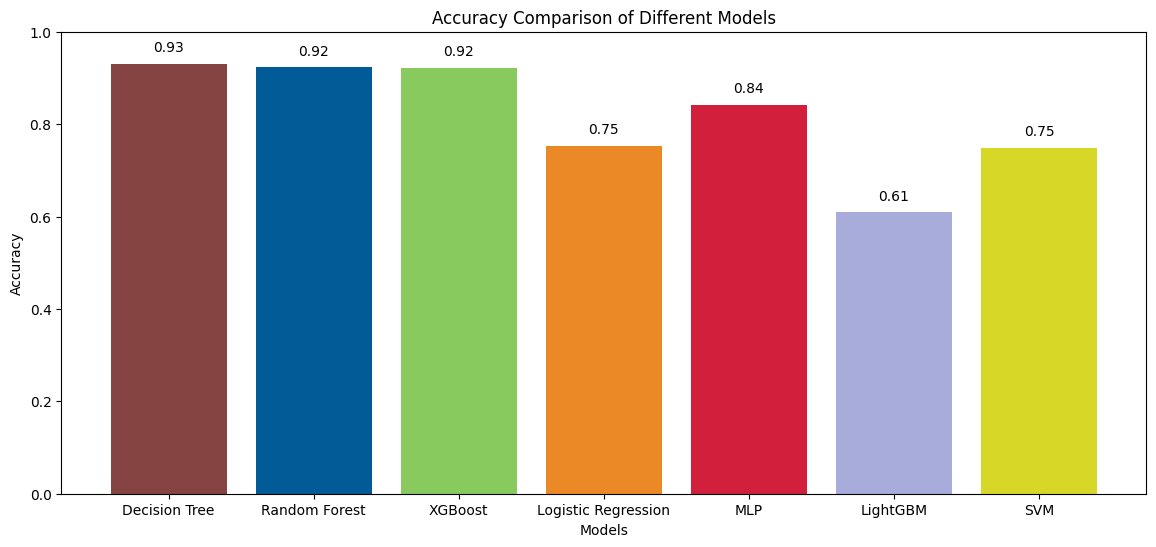


Figure: Comparison of different Machine Learning models accuracy.

**4.3 Deep Learning Models**

In this segment, we dive into the domain of deep learning, assessing the viability of two unmistakable structures, Artificial Neural Networks (ANN) and one-dimensional Convolutional Neural Networks (1D CNN), with regards to anticipating thyroid illness. These models bridle the force of neural networks to gain perplexing examples and portrayals from the information, offering a promising road for medical diagnostics consequently. Table 02 Shows the consequences of various assessment measurements of Deep Learning models.

**Table 02**: Result analysis of Deep Learning models

|  |  |  |  |
| --- | --- | --- | --- |
| **Model Name** | **Accuracy** | **Precision** | **F1 Score** |
| ANN | 96% | 94% | 87% |
| 1D CNN | 90% | 81% | 85% |
| DNN | 78.20% | 68% | 72% |
| Hybrid Model | 75% | 74% | 74% |

The ANN model exhibited a noteworthy accuracy pace of 0.96 and effectively anticipated the two classes with high precision and recall. Then again, the 1D CNN model performed well in foreseeing class 0 with an accuracy pace of 0.90 and high precision and recall, yet battled in anticipating class 1, bringing about lower precision, recall, and F1-score. Contrasting the performance of deep learning models versus top-performing machine learning models, we saw that while the ANN model showed serious accuracy, precision, and recall, the 1D CNN model succeeded in catching genuine positive cases for class 0 yet experienced hardships in anticipating class 1 precisely. These outcomes exhibit the capability of deep learning models, especially ANNs, in accomplishing high prescient accuracy in thyroid disease prediction. Their capacity to naturally gain complex examples from information features their significance in medical diagnostics. Be that as it may, difficulties, for example, class imbalances and model architecture limitations should be addressed to understand their potential completely. All in all, deep learning models, particularly ANNs, offer promising roads for working on analytic accuracy in thyroid disease prediction. Further examination and streamlining are important to address the limitations noticed, preparing for the incorporation of deep learning procedures into clinical practice and enhancing patient results.

**4.4 LIME Explainable AI (XAI)**

We employed LIME to analyze the outcomes of two machine learning models - Random Forest XGBoost and Decision Tree. Our goal was to use Explainable Artificial Intelligence (XAI) to identify the most important features for the models to comprehend the predictions of all black box models. This gave us explanations at the local level for a specific instance. The Lime explanation for Random Forest XGBoost and Decision Tree models can be seen in Figures 7, 8, and 9 respectively.

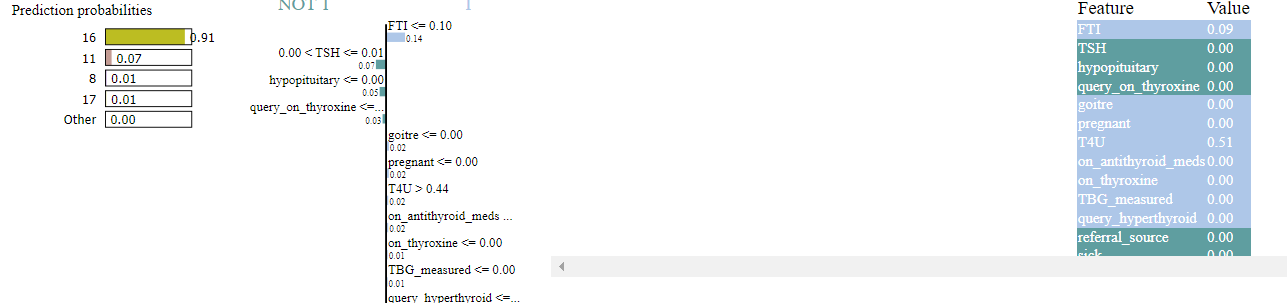
Figure 7: Lime explanation for Random Forest

Figure 07 illustrates the LIME explanation for the Random Forest Model, which scrutinizes a sample row to predict the probability of patient status. The prediction assigns class number sixteen (16), denoting 'psych' for the selected row. 'Psych' indicates the whether patient is psych. Notably, the features with the most significant impact on this prediction are 'TSH,' 'hypopituitary,' and 'query\_on\_thyroxine.'

A close-up of a person

Description automatically generated

Figure 8: Lime explanation for Decision Tree

Figure 08 depicts the LIME explanation for the Decision Tree Model, which examines a sample row to predict the probability of patient status. The prediction assigns class number (03), corresponding to 'on\_thyroxine' for the selected row. 'On\_thyroxine' indicates whether the patient is on thyroxine medication. Importantly, the features with the most significant impact on this prediction are 'sex,' 'TSH\_measured,' and 'TT4.'

A screenshot of a computer

Description automatically generated

Figure 9: Lime explanation for XGBoost

Figure 09 illustrates the LIME explanation for the XGBoost Model, which analyzes a sample row to predict the probability of patient status. The prediction assigns class number (03), corresponding to 'on\_thyroxine' for the selected row. 'On\_thyroxine' indicates whether the patient is on thyroxine medication. Importantly, the features with the most significant impact on this prediction are 'FSH,' 'T3,' and 'goitre.'.

**Chapter 5**

**Conclusion & Future Works**

## 5.1 Conclusion

All in all, our review has shown the viability of different machine learning and deep learning models in foreseeing thyroid disease. Through thorough assessment utilizing standard measurements like accuracy, precision, recall, and F1-score, we have gained important experiences into the performance of these models across various prescient errands.

From our analysis, it is obvious that troupe techniques like Random Forest and XGBoost, close by interpretable models, for example, Decision Trees, display unrivaled performance in catching fundamental examples inside the dataset. These models show robustness in speculation and exact prediction, making them promising candidates for thyroid disease prediction assignments. Furthermore, our While the ANN model showed competitive performance across all metrics, the 1D exploration into deep learning techniques, particularly Artificial Neural Networks (ANN) and one-dimensional Convolutional Neural Networks (1D CNN), highlights their potential in achieving high predictive accuracy. CNN model excelled in capturing specific patterns but encountered challenges in predicting certain classes accurately.

Moreover, the application of Explainable Artificial Intelligence (XAI) through LIME provided valuable insights into the inner workings of the models, allowing us to interpret their predictions and understand the significance of different features in thyroid disease diagnosis.

## 5.2 Future Works

Pushing ahead, there are a few roads for future innovative work in the field of thyroid illness expectation: Further streamlining of AI and deep learning models, including hyperparameter tuning and design changes, could upgrade prescient execution and speculation capacities. Investigation of extra highlights and refinement of existing component portrayals might work on the discriminative force of the models and uncover new experiences into thyroid sickness determination. Usage of information expansion methods to address class uneven characters and increment the variety of the dataset could assist with moderating predispositions and work on the strength of prescient models. Examination concerning novel troupe strategies and crossover draws near, joining the qualities of various models, may prompt further enhancements in prescient accuracy and dependability. Coordinated effort with medical services experts and reconciliation of prescient models into clinical decision support frameworks could work with genuine applications and upgrade patient results in thyroid sickness the board. Proceeded with examination into Reasonable computer based intelligence procedures, including the improvement of additional interpretable models and the refinement of existing strategies like LIME, can upgrade our understanding of model expectations and cultivate trust in computer based intelligence driven clinical diagnostics. By tending to these future works, we can propel the field of thyroid infection expectation, eventually prompting more precise, solid, and interpretable models that benefit the two patients and medical services professionals.

# 

# References

[1] L. Ozyılmaz and T. Yıldırım, "Diagnosis of thyroid disease using artificial neural network methods," in Proceedings of ICONIP’02 9th International Conference on Neural Information Processing (Singapore: Orchid Country Club, 2002), pp. 2033–2036.

[2] K. Polat, S. Sahan, and S. Gunes, "A novel hybrid method based on artificial immune recognition system (AIRS) with fuzzy weighted preprocessing for thyroid disease diagnosis," Expert Systems with Applications, vol. 32, 2007, pp. 1141-1147.

[3] F. Saiti, A. A. Naini, M. A. Shoorehdeli, and M. Teshnehlab, "Thyroid Disease Diagnosis Based on Genetic Algorithms Using PNN and SVM," in 3rd International Conference on Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009.

[4] G. Zhang and L.V. Berardi, "An investigation of neural networks in thyroid function diagnosis," Health Care Management Science, 1998, pp. 29-37. Available: http://www.endocrineweb.com/thyroid.html, (Accessed: 7 August 2007).

[5] V. Vapnik, Estimation of Dependences Based on Empirical Data, Springer, New York, 2012.

[6] Obermeyer Z. and Emanuel E.J., "Predicting the future— big data, machine learning, and clinical medicine," N Engl J Med, 2016; 375:1216-1219.

[7] Breiman L., "Statistical Modeling: the two cultures," Stat Sci, 2001;16:199-231.

[8] Ehrenstein V., Nielsen H., Pedersen AB., Johnsen SP., and Pedersen L., "Clinical epidemiology in the era of big data: new opportunities, familiar challenges," Clin Epidemiol, 2017; 9:245-250.

[9] Ghahramani Z., "Probabilistic machine learning and artificial intelligence," Nature, 2015;521: 452-459.

[10] Azimi P., Mohammadi HR., Benzel EC., Shahzadi S., and Azhari S., "Artificial neural networks in neurosurgery," J Neurol Neurosurg Psychiatry, 2015; 86:251-256.

[11] Deo RC., "Machine learning in medicine," Circulation, 2015;132:1920-1930.

[12] P.C. Austin, J.V. Tu, J.E. Ho, D. Levy, and D.S. Lee, "Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes," J Clin Epidemiol, 2013;66(4):398–407.

[13] A.K. Pandey, P. Pandey, and K.L. Jaiswal, "A heart disease prediction model using Decision Tree," IUP J Comput Sci, 2013;7(3):43.

[14] S. Ismaeel, A. Miri, and D. Chourishi, "Using the Extreme Learning Machine (ELM) technique for heart disease diagnosis," IEEE Canada International Humanitarian Technology Conference, 2015, pp. 1–3.

[15] L. Verma, S. Srivastava, and P.C. Negi, "A hybrid data mining model to predict coronary artery disease cases using noninvasive clinical data," J Med Syst, 2016;40(7):1–7.

[16] R. Rajkumar, K. Anandakumar, and A. Bharathi, "Coronary artery disease (CAD) prediction and classification—a survey," ARPN J Eng Appl Sci, 2006;11(9):5749–5754.

[17] Y.T. Lo, H. Fujita, and T.W. Pai, "Prediction of coronary artery disease based on ensemble learning approaches and co-expressed observations," J Mech Med Biol, 2016;16(01):1640010.

[18] B. Farran, A.M. Channanath, K. Behbehani, and T.A. Thanaraj, "Predictive models to assess risk of type 2 diabetes, hypertension and comorbidity: machine-learning algorithms and validation using national health data from Kuwait—a cohort study," BMJ Open, 2013;3(5):e002457.

[19] M. Heydari, M. Teimouri, Z. Heshmati, and S.M. Alavinia, "Comparison of various classification algorithms in the diagnosis of type 2 diabetes in Iran," Int J Diabetes Dev Countries, 2015;36(2):167–173.

[20] Prerana, Parveen Sehgal, Khushboo Taneja, "Predictive Data Mining for Diagnosis of Thyroid Disease using Neural Network," International Journal of Research in Management, Science & Technology, Vol. 3, No. 2, April 2015.

[21] S. Sathya Priya, Dr. D. Anitha, "Survey on Thyroid Diagnosis using Data Mining Techniques," International Journal of Advanced Research in Computer and Communication Engineering, Vol. 6, Special Issue 1, January 2017.

[23] Wajner, S.M. and Maia, A.L., "New insights toward the acute non-thyroidal illness syndrome," Front. Endocrinol., 2012, 3:8.

[24] Cai, J., Luo, J., Wang, S., and Yang, S., "Feature selection in machine learning: A new perspective," Neurocomputing, 2018, 300, 70–79.

[25] Solorio-Fernández, S., Carrasco-Ochoa, J.A., and Martínez-Trinidad, J.F., "A review of unsupervised feature selection methods," Artif. Intell. Rev., 2020, 53, 907–948.

[26]Tech,G.,"FeatureSelection,"Availableonline:<https://faculty.cc.gatech.edu/~bboots3/CS4641-Fall2018/Lecture16/16_FeatureSelection.pdf> (accessed on 7 May 2022).

[27] Baby, D., Devaraj, S.J., Hemanth, J., and Anishin Raj, M.M., "Leukocyte classification based on feature selection using extra trees classifier: A transfer learning approach," Turk. J. Electr. Eng. Comput. Sci., 2021, 29, 2742–2757.

[28] Leng, L. and Zhang, J., "Palmhash code vs. palmphasor code," Neurocomputing, 2013, 108, 1–12.

[29] Rustam, F., Ishaq, A., Munir, K., Almutairi, M., Aslam, N., and Ashraf, I., "Incorporating CNN Features for Optimizing Performance of Ensemble Classifier for Cardiovascular Disease Prediction," Diagnostics, 2022, 12, 1474.