



**TITLE**

**PREDICTION OF THE LEVEL OF HYPOTHYROIDISM HORMONES FOR THE MIDDLE AGE GROUP AND PROTOTYPE MODEL OF MODIFIED PUNCH BIOPSY**

**A COURSE LEVEL PROJECT REPORT**

***Submitted by***

**III-year students of Bachelor of Technology**

R. SESHU VARDHAN REDDY-

99210042076

R. SHIRI SAI CHANDANA-

99210041626

G. SHIVA CHAITANAYA-

99210042047

G. SRINIVAS SIDDHARTH-

9921004237

V. SUJAY SAI REDDY-

9921004768

***in partial fulfillment of the course of***

**215EXS3201 / EXSEL – DESIGN-BUILD-OPERATE**

**Academic Year 2023 – 2024 (Even Semester)**



**BONAFIDE CERTIFICATE**

Certified that this project report **“Prediction of the level of hypothyroidism hormones for the middle age group and prototype model of modified punch biopsy”** is the bonafide work of **“…………………………….”** who carried out the project work under my supervision.

**Faculty Mentor**

Submitted for the Project Viva-voce / Review held at Kalasalingam Academy of Research & Education, Krishnankoil on ………………………………

**Internal Examiner External Examiner**

**PROBLEM STATEMENT**

* To implement precision Tissue sampling to ensure accuracy in obtaining high-quality punch biopsy.
* Prediction of the level of hypothyroidism hormones for the middle age group using a dataset containing age, gender, thyroid-stimulating hormone (TSH) levels, triiodothyronine (T3), and thyroxine (T4).
* Designing a diet chart for middle-aged group people using Python.

**PS No.:** WB-PS30

**Team No.:** WB41

**Abstract**

Hypothyroidism is a common endocrine disorder that affects millions of people worldwide, with a higher prevalence in middle-aged individuals. Accurate prediction of hypothyroidism hormone levels is crucial for timely diagnosis and treatment. This study aims to develop 5 machine learning-based models Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient boosting on a dataset related to hypothyroidism prediction. To predict hypothyroidism hormone levels in middle-aged individuals and design a prototype of a modified punch biopsy technique for thyroid tissue sampling.

Predicting the level of hypothyroidism hormones in middle-aged individuals can be approached through various methods, including machine learning algorithms trained on relevant medical data. These algorithms can analyze factors such as age, gender, thyroid-stimulating hormone (TSH) levels, triiodothyronine (T3), thyroxine (T4), and other pertinent medical history to make predictions.

In parallel, we designed a modified punch biopsy prototype incorporating a novel needle design and sampling protocol to minimize tissue damage and improve sample quality. The integration of these two approaches can facilitate early diagnosis and treatment of hypothyroidism, ultimately improving patient outcomes.

**Introduction**

Punch biopsy is a minimally invasive procedure for collecting thyroid tissue samples for histopathological examination. However, traditional punch biopsy techniques can be associated with bleeding complications, tissue damage, and inadequate sample quality. The development of a proposed punch biopsy technique that can improve tissue yield and reduce complications would be a significant advancement in the diagnosis and management of hypothyroidism.

Machine learning algorithms Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient Boosting a dataset related to hypothyroidism prediction have been increasingly used in medicine to predict disease outcomes, diagnose conditions, and identify high-risk patients. These algorithms can analyze large datasets, recognize complex patterns, and make accurate predictions. In the context of hypothyroidism, machine learning algorithms can be used to predict hormone levels based on clinical and laboratory data, enabling early diagnosis and treatment.

Methods for detecting thyroid levels typically include blood tests and imaging tests like ultrasounds or scans, as well as a physical examination by a healthcare provider. Common blood tests, known as Thyroid Function Tests (TFTs), include the Thyroid-Stimulating Hormone (TSH) Test, which measures the level of TSH in the blood, stimulating the thyroid gland to produce thyroid hormones. The Total Thyroxine (TT4) Test measures the total amount of T4 in the blood, including both bound and free forms, while the Total Triiodothyronine (TT3) Test measures the total amount of T3 in the blood, also including both bound and free forms. Another laboratory test used is the Enzyme-Linked Immunosorbent Assay (ELISA), which measures the levels of thyroid hormones in the blood using enzymes and antibodies. Imaging tests are also essential in thyroid diagnostics; an ultrasound uses high-frequency sound waves to produce images of the thyroid gland. The Radioactive Iodine Uptake (RAIU) Test measures the uptake of radioactive iodine by the thyroid gland, aiding in the diagnosis of thyroid disorders, and Thyroid Scintigraphy uses small amounts of radioactive material to produce images of the thyroid gland.

A biopsy, a surgical procedure where a sample of thyroid tissue is removed for microscopic examination, can provide further insight. The TSH Test is particularly crucial as it measures the level of TSH produced by the pituitary gland to stimulate the thyroid gland, with elevated TSH levels typically indicating hypothyroidism. Additionally, tests measuring Thyroxine (T4) and Triiodothyronine (T3) levels assess the two primary thyroid hormones, where low levels of T4 and T3 indicate potential thyroid dysfunction.

Detecting thyroid levels is crucial for diagnosing and managing thyroid disorders, as it helps in understanding the functioning of the thyroid gland and provides insights into potential health risks. Early detection can prevent complications and promote better management of thyroid conditions, enabling timely and effective interventions. Regular monitoring of thyroid levels allows healthcare providers to make informed treatment decisions, ensuring that patients receive the most appropriate care based on their thyroid hormone levels. Identifying imbalances in thyroid levels is essential for restoring optimal thyroid function through appropriate interventions, thereby maintaining overall health and well-being.

Key nutrients essential for thyroid health include iodine and selenium. Iodine is crucial for thyroid hormone production, while selenium plays a vital role in protecting the thyroid gland from oxidative stress. Good sources of iodine are seafood, dairy products, and iodized salt. Selenium can be found in Brazil nuts, sunflower seeds, and whole grains. By comparing 5 algorithms- Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient Boosting a dataset related to hypothyroidism prediction and find the best algorithm and its accuracy.

**Literature Survey**

Wen-Juan Li et al [1] proposed thyroid disease, thyroid surgery, or radioactive iodine therapy will be excluded from the study. The data collection process will involve gathering demographic data, medical history, and laboratory results, including TSH, FT4, and FT3 levels. The study will have a 2-year follow-up period, during which the participants will be observed for any changes in their health status. Limited generalizability study is limited to elderly individuals in a specific region, which may not be representative of other populations.2023

Sosa, J. A. et al [2] it is a Comprehensive literature review of published studies on thyroid surgery, including randomized controlled trials, observational studies, and case series. Consensus-based approach, where the panel members discussed and agreed upon the recommendations. The grading system, where the strength of the recommendations was based on the quality of the evidence. The guidelines are based on a consensus-based approach, which may be subjective and influenced by individual biases.2020

Yang, J. et al [3] have studied how to evaluate the effects of thyroid FNAB on serum thyroid hormone levels, including total triiodothyronine (TT3), total thyroxine (TT4), free triiodothyronine (FT3), free thyroxine (FT4), and thyroid-stimulating hormone (TSH). Data extraction Two reviewers independently extracted data from the included studies, including study characteristics, patient demographics, and serum thyroid hormone levels before and after FNAB. The study did not perform a meta-analysis, which could have provided a more quantitative synthesis of the results.2019

Huh, J. et al [4] examined the 50 patients with thyroid nodules who underwent FNAB and were enrolled in the study. Serum thyroid hormone levels, including total triiodothyronine (TT3), total thyroxine (TT4), free triiodothyronine (FT3), free thyroxine (FT4), and thyroid-stimulating hormone (TSH), were measured before and 1 week after FNAB. It had a small sample size, which may not be representative of the general population.2014

Fugazzola, L. et al [5] have studied Serum thyroid-related constituents, including thyroid-stimulating hormone (TSH), free thyroxine (FT4), free triiodothyronine (FT3), and thyroglobulin (Tg), were measured before and 1, 3, and 7 days after FNAB. The patients were divided into two groups based on the presence or absence of thyroid autoantibodies (TAb). And proposed a small sample size, which may not be representative of the general population.2010

Fugazzola, L. et al [6] have proposed and searched for thyroid fine-needle biopsy and serum markers. Selection Criteria They would have defined specific inclusion and exclusion criteria for selecting studies, possibly focusing on those that reported changes in serum thyroid-related constituents post-FNB. Study Heterogeneity Variability in study designs, patient populations, biopsy techniques, and assays used for serum markers can introduce heterogeneity that may limit direct comparisons or the generalizability of findings.2010

Lily L. et al [7] they have designed the cohort design, possibly retrospective or prospective, to examine associations between subclinical thyroid disease markers (like TSH and thyroid hormones) and cardiovascular events or mortality. **Causality** like many observational studies, establishing causality between subclinical thyroid disease and cardiovascular outcomes can be challenging due to the potential for unmeasured confounders.

**Survey on existing products**

A punch biopsy is a medical procedure used to obtain a small tissue sample from the body for diagnostic purposes. This technique is commonly employed in dermatology but can also be used in other medical specialties. The procedure involves using a specialized instrument called a biopsy punch, which is a circular blade or cylindrical tool that is used to remove a cylindrical core of tissue from the targeted area.

**General overview of the punch biopsy procedure:**

Preparation: The patient's skin or the targeted area is cleaned and sterilized to reduce the risk of infection.

**Local Anesthesia:** A local anesthetic is typically administered to numb the area where the biopsy will be performed. This helps minimize discomfort during the procedure.

**Biopsy Punch:** The biopsy punch, which comes in various sizes, is selected based on the lesion size or area of interest. The healthcare professional places the punch over the target site and rotates it to cut through the layers of skin and collect a tissue sample.

**Existing Methodology**

**Tissue Collection:** The cylindrical core of tissue obtained with the biopsy punch is then removed. In some cases, stitches may be required to close the wound, especially if a larger sample is taken.

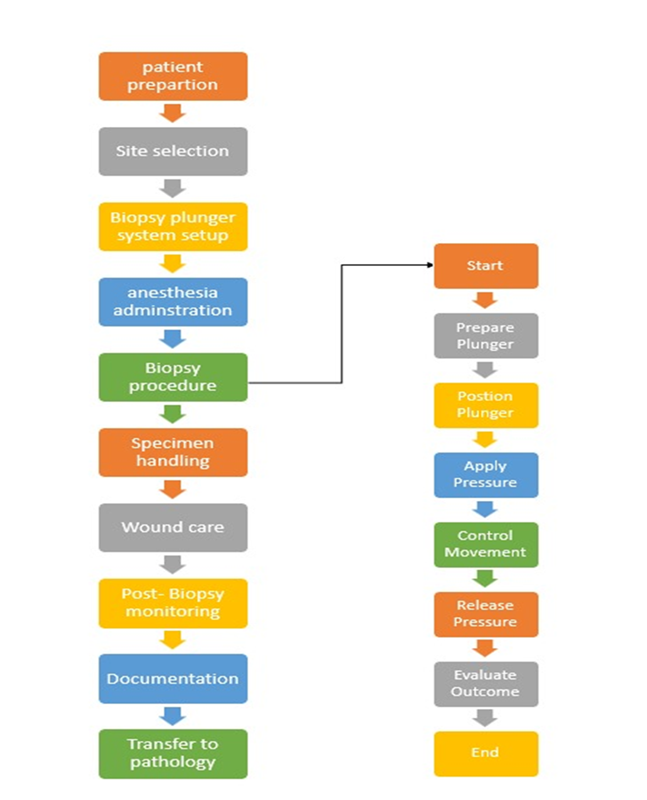
**Wound Care:** After the procedure, the wound is typically covered with a sterile dressing. The patient is given instructions on how to care for the wound to promote proper healing and minimize the risk of infection.

Punch biopsies are often preferred for skin lesions or abnormalities because they are relatively quick, simple, and associated with minimal scarring. However, the specific type of biopsy used may depend on the characteristics of the lesion, the suspected diagnosis, and the location of the body. As with any medical procedure, there are potential risks and complications, and these should be discussed with the healthcare provider before the biopsy is performed.

**Proposed**

We propose a new punch biopsy prototype aimed at improving diagnostic accuracy and patient comfort. Additionally, we are working on a project for predicting hypothyroidism in middle-aged groups using Gradient Boosting, a powerful machine-learning technique. Moreover, we are designing a diet chart tailored for middle-aged individuals utilizing Python, ensuring nutritional balance and health optimization.

**Flowchart:**



- The punch biopsy comes with a user-friendly, one-time-use plunger system design.

**Material Selection:**

- Our product uses cost-effective plastic rather than expensive steel due to its one-time-use nature.

Types of MM Blades:1.5MM,2MM,3MM,3.5MM,4MM,5MM,6MM,8MM,10MM

**Prototyping:**

- We transformed the punch biopsy design into a click pen model, a plunger system.

**Modification:**

- We inserted a spring into the body of the punch biopsy to facilitate the easy movement of the mm blade, making it easier to use.

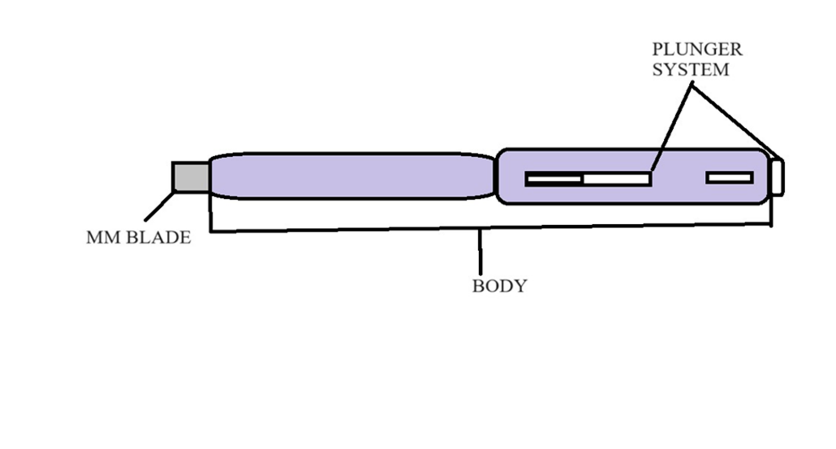
**Ergonomic Designs:**

We are considering improving the biopsy instruments to enhance ergonomics for healthcare professionals, ensuring a more comfortable and efficient procedure.

**Biopsy Plunger Systems:**

While the "punch biopsy plunger system" may not be widely known, it is a crucial component in our design. Recognized, it may refer to a specific type of instrument or system designed to improve the ease and efficiency of obtaining tissue samples during a punch biopsy. This could include mechanisms for controlled sample retrieval.

**Proposed Punch Biopsy Design**



**Diet Chart**

The code starts by defining four lists to store the diet chart information: breakfast, a list of healthy breakfast options; lunch, a list of healthy lunch options; dinner, a list of healthy dinner options; and snacks, a list of healthy snack options. Additionally, it includes avoid\_foods, a list of foods to avoid or limit for people with thyroid problems.

To create a pie chart representing the distribution of foods in the diet chart, the code defines two lists: categories and percentages. The categories list includes "Low iodine foods" (60%), "Moderate iodine foods" (25%), "High iodine foods" (10%), and "Foods to avoid" (5%). Corresponding percentages are stored in the percentages list.

Colors for the pie chart are specified in a list, where "#ff9900" (orange) is used for "Low iodine foods," "#00ffcc" (teal) for "Moderate iodine foods," "#66b3ff" (blue) for "High iodine foods," and "#c2c2f0" (gray) for "Foods to avoid."

The code utilizes the plot. Pie function to create the pie chart, passing the percentages, categories, and color lists as arguments. The autopilot argument is set to "%.1f%%" to display the percentage value for each category on the chart.

**Experimental result**

**Hypothyroidism Prediction Analysis:**

The goal is to predict whether an individual has hypothyroidism based on features like age, gender, and thyroid hormone levels (T3, T4, TSH). The prediction uses a logistic regression model, which is well-suited for binary classification problems.

The dataset used in this analysis, provided by the medical faculty at Kalasalingam University, aims to predict hypothyroidism among individuals. It includes 89 entries, with 72 females and 17 males. Key attributes in the dataset are age, gender, T3 (triiodothyronine) level, T4 (thyroxine) level, TSH (thyroid-stimulating hormone) level, and a binary indicator for hypothyroidism (HYP), where 1 indicates the presence of hypothyroidism and 0 indicates its absence. The data is loaded from a CSV file and preprocessed to handle missing values by replacing empty strings with NaN and subsequently dropping any rows containing NaNs. Categorical variables, such as gender, are encoded into numerical values using LabelEncoder. For feature selection, the predictors include age, gender, T3, T4, and TSH levels, while the target variable is hypothyroidism status.

To focus on a specific subset, the dataset is filtered to include only middle-aged individuals (ages 35-55) who have hypothyroidism. The data is then split into training and testing sets, with 80% used for training and 20% for testing. Features are scaled using StandardScaler to normalize the data. A logistic regression model is trained on the training data and evaluated on the testing data using accuracy and a confusion matrix. Additionally, visualization is performed to illustrate the distribution of thyroid hormone levels (T3, T4, TSH) among middle-aged individuals with hypothyroidism, with histograms plotted separately for males and females. The logistic regression algorithm, a statistical method for analyzing datasets with one or more independent variables that determine a binary outcome, is the primary algorithm used in this analysis. The outcome of this method indicates whether an individual has hypothyroidism (1) or not (0).

The performance of five different machine learning algorithms-Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient boosting on a dataset related to hypothyroidism prediction. The dataset is preprocessed by replacing empty strings with NaN values, dropping rows with missing data, and encoding the 'GENDER' column. The features used for prediction include 'AGE', 'GENDER', 'T3', 'T4', and 'TSH', while the target variable is 'HYP', which indicates the presence of hypothyroidism. After splitting the data into training and testing sets and scaling the features, each model is trained and evaluated. The script calculates the accuracy and confusion matrix for each model, which is then printed for comparison. The model with the highest accuracy is identified as the best model.

**The comparison results showed the following accuracies for each model:**

Logistic Regression: Achieved an accuracy of 0.6666666666666666.

Decision Tree: Achieved an accuracy of 0.6111111111111112.

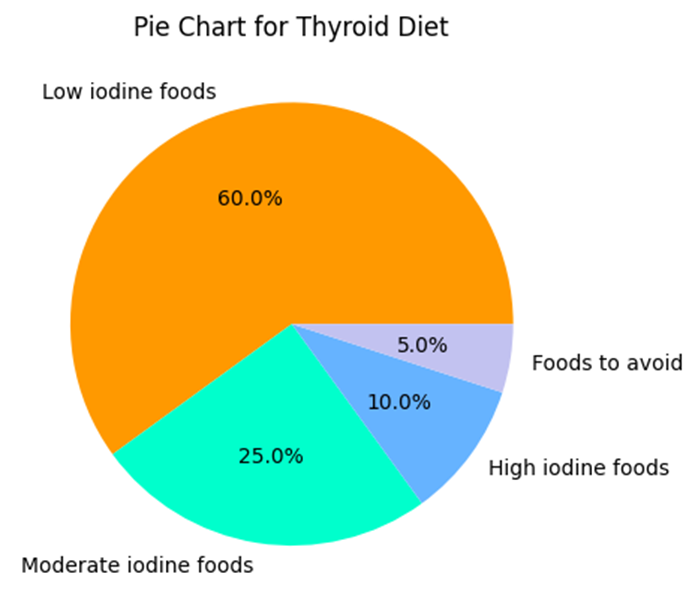
Random Forest: Achieved an accuracy of 0.5555555555555556.

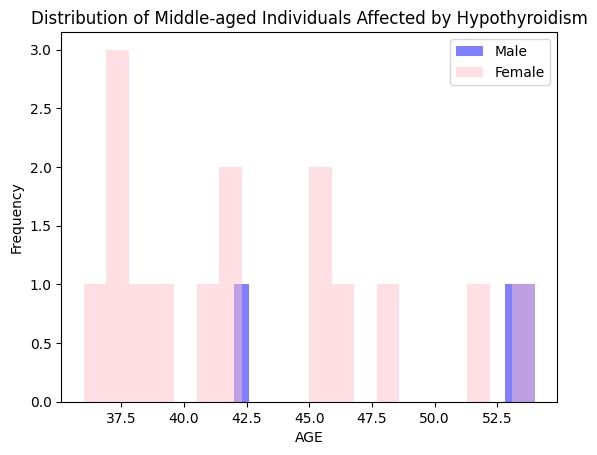
Support Vector Machine (SVM): Achieved an accuracy of 0.6666666666666666.

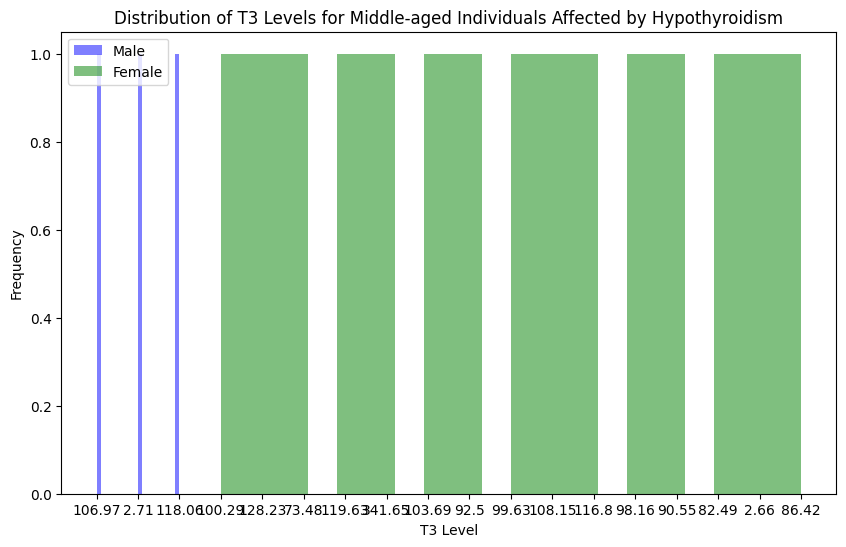
Gradient Boosting: Achieved an accuracy of 0.7222222222222222.

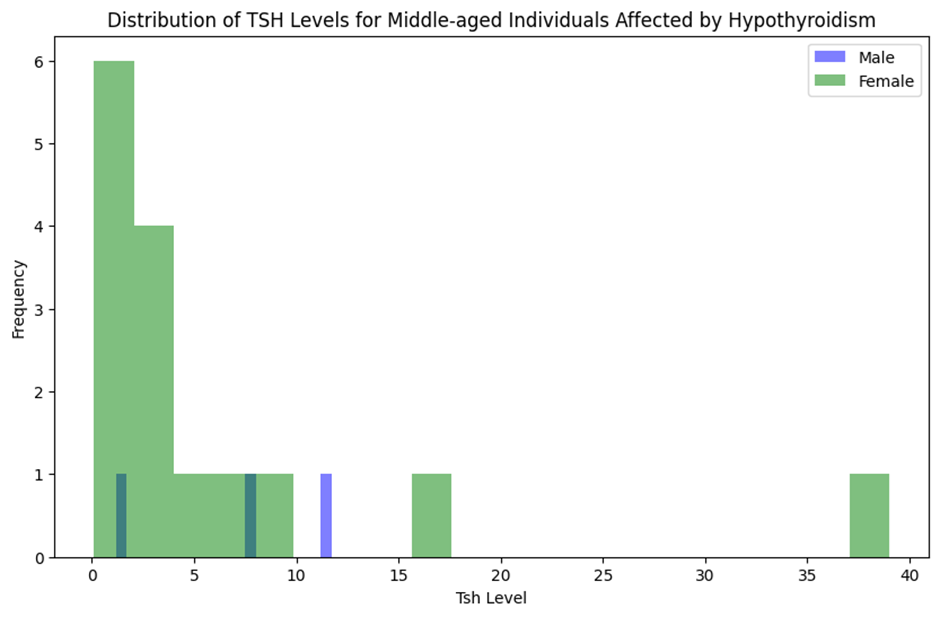
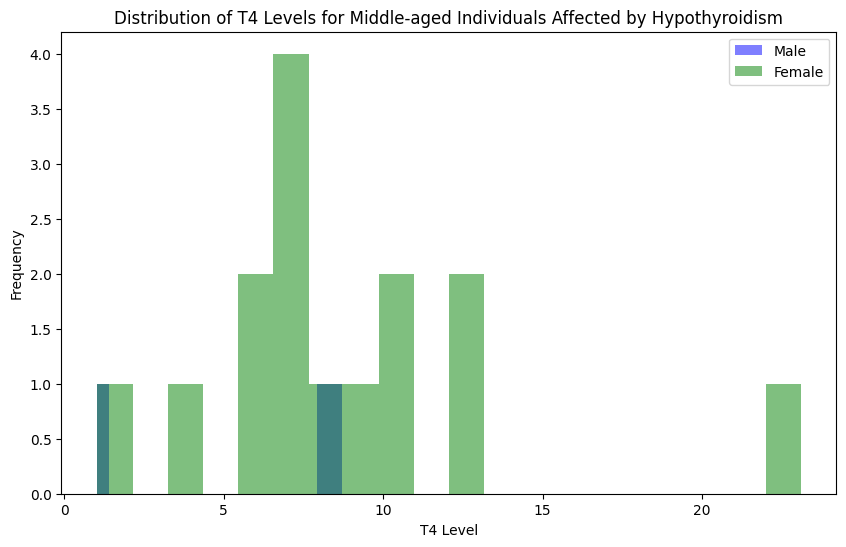
Overall, the Gradient Boosting model emerged as the best-performing algorithm with the highest accuracy of 0.7222222222222222. This suggests that Gradient Boosting is the most effective algorithm among those tested for predicting hypothyroidism in this dataset. The subsequent analysis and visualizations focused on middle-aged individuals (ages 35-55) with hypothyroidism, highlighting the distribution of T3, T4, and TSH levels for males and females, further aiding in understanding the condition's demographic characteristics.

**Outcomes**

****







**Conclusion**

For the detection of thyroid disorders, a comprehensive approach involves clinical assessments, laboratory tests, imaging studies, and potentially leveraging advanced technologies like machine learning and AI algorithms. This methodology considers the complexity of thyroid conditions, incorporates a range of diagnostic tools, and emphasizes the importance of continuous monitoring and collaboration with specialists. The double-click pen model of modified punch biopsy offers several advantages over traditional biopsy techniques. Its customizable depth control allows for tailored sampling of superficial or deep lesions, ensuring optimal tissue retrieval for diagnostic evaluation. Additionally, the ergonomic design of the double-click pen enhances usability and maneuverability, facilitating precise positioning and execution of biopsies even in challenging anatomical locations.

Furthermore, the double-click pen model promotes procedural efficiency and streamlines workflow in clinical settings. Its intuitive operation and rapid adjustment capabilities enable clinicians to perform biopsies with increased speed and accuracy, ultimately reducing procedural time and enhancing patient satisfaction. Moreover, the device's compatibility with standard biopsy needles ensures versatility and ease of integration into existing biopsy protocols and practices. In conclusion, the double-click pen model of modified punch biopsy represents a significant technological innovation in dermatological diagnostics, offering clinicians a versatile and efficient tool for obtaining high-quality skin tissue samples. With its precision, usability, and procedural benefits, the double-click pen model holds promise for advancing the field of dermatopathology and improving patient care outcomes.

Incorporating advanced technologies such as machine learning Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient Boosting a dataset related to hypothyroidism prediction. And AI algorithms in the comprehensive approach for thyroid disorder detection further enhance diagnostic accuracy. These algorithms can analyze a range of factors, including age, gender, thyroid-stimulating hormone (TSH) levels, triiodothyronine (T3), thyroxine (T4), and other pertinent medical history, to make precise predictions.

Designing a diet chart for middle-aged individuals involves careful consideration of nutritional needs to support overall health and well-being. The focus should be on achieving a balanced intake of macronutrients- carbohydrates, proteins, and health through foods like whole grains, lean proteins such as fish and legumes, and sources of unsaturated fats like nuts and olive oil. It's crucial to emphasize a variety of colorful fruits and vegetables rich in vitamins, minerals, and antioxidants, aiming for at least 5 servings daily. Limiting processed foods and sugars helps manage weight and reduces the risk of chronic diseases. Additionally, ensuring adequate hydration, portion control, and considering individual preferences and health goals are essential for creating a sustainable and effective diet plan tailored to middle-aged adults. By comparing 5 algorithms- Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Gradient Boosting a dataset related to hypothyroidism prediction then the best machine learning model is Gradient Boosting.

**Reference**

[1] Wen-Juan Li, Xiao-Li Wang, Li-Hao Sun, Qian-Qian Nie, Hua-Hua Liu, and Rong-Hua Song.The different outcomes in the elderly with subclinical hypothyroidism diagnosed by age-specific and non-age-specific TSH reference intervals: a prospectively observational study protocol. Clinical and Experimental Thyroidology.2023

[2] Sosa, J. A., Hanna, J. W., Robinson, K. A., & Lanman, R. B. The American Association of Endocrine Surgeons Guidelines for the Definitive Surgical Management of Thyroid Disease in Adults. Annals of Surgery, 271(3), e21-e93.2020

[3] Yang, J., Sun, X., Shen, J., & Li, W. Effects of Thyroid Fine-Needle Aspiration Biopsy on Serum Thyroid Hormone Levels: A Systematic Review. International Journal of Endocrinology, 2019, Article ID 6212759.2019

[4] Huh, J. Y., Park, S. Y., Cho, M. S., & Bae, J. S. Effects of Thyroid Fine-Needle Aspiration Biopsy on Serum Thyroid Hormone Levels. Endocrinology and Metabolism, 29(4), 481-486.2014

[5] Fugazzola, L., Mannavola, D., Cirello, V., Beck-Peccoz, P., & Chiovato, L. Alterations in Serum Thyroid-Related Constituents After Thyroid Fine-Needle Biopsy. Thyroid, 20(2),121-127.2010

[6] Fugazzola, L., Mannavola, D., Cirello, V., Beck-Peccoz, P., & Chiovato, L. Alterations in Serum Thyroid-Related Constituents After Thyroid Fine-Needle Biopsy. Thyroid, 20(2), 197-203. 2010

[7] Lily L. Somwaru, Alice M. Arnold, Neha Joshi, Linda P. Fried, Anne R. Cappola, Association of Subclinical Thyroid Disease with Cardiovascular Diseases and Mortality in the Elderly Clinical Endocrinology and Metabolism (JCEM) in, Volume 94, Issue 4, 2009, Pages 1342–1348.2009