THYROID DISEASE CLASSIFICATION USING SUPPORT VECTOR MACHINE AND RECURSIVE FEATURE ELIMINATION METHOD

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh
CITRA WULANDARI
NIM. 12150121717

UIN SUSKA RIAU

FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2026

LEMBAR PERSETUJUAN

THYROID DISEASE CLASSIFICTION USING SUPPORT VECTOR MACHINE AND RECURSIVE FEATURE ELIMINATION METHOD

LAPORAN TUGAS AKHIR MAHASISWA
JURUSAN TEKNIK INFORMATIKA
UIN SUSKA RIAU

TUGAS AKHIR

Oleh

CITRA WULANDARI
NIM. 12150121717

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 7 Januari 2026

Pembimbing I

IIS AFRIANTY, S.T., M.Sc
NIP. 19880426201903 2 009

LEMBAR PENGESAHAN

THYROID DISEASE CLASSIFICTION USING SUPPORT VECTOR MACHINE AND RECURSIVE FEATURE ELIMINATION METHOD

LAPORAN TUGAS AKHIR MAHASISWA
JURUSAN TEKNIK INFORMATIKA
UIN SUSKA RIAU

Oleh

CITRA WULANDARI
NIM. 12150121717

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 7 Januari 2026

Mengesahkan,

Dekan,	Ketua Jurusan,
Dr. Yuslenita Muda, S.Si., M.Sc.	**Muhammad Affandes, S.T., M.T.**
NIP. 19770103 200710 2 001	**NIP. 19861206 201503 1 004**

DEWAN PENGUJI

Ketua	: Pizaini, S.T., M.Kom.
Pembimbing I	: Iis Afrianty, S.T., M.Sc
Penguji I	: Elvia Budianita, S.T., M.Cs.
Penguji II	: Siska Kurnia Gusti, S.T., M.Sc

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama	: Citra Wulandari
NIM	: 12150121717
Tempat/Tgl Lahir	: Pekanbaru, 29 Agustus 2002
Fakultas	: Sains dan Teknologi
Prodi	: Teknik Informatika
Judul Skripsi	: Thyroid Disease Classification Using Support Vector Machine and Recursive Feature Elimination Method

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan jurnal dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis ini sudah disebutkan sumbernya.
3. Oleh karena itu jurnal saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan jurnal saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikian surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 12 Januari 2026

Yang membuat pernyataan

CITRA WULANDARI

NIM.12150121717

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul "Thyroid Disease Classification Using Support Vector Machine and Recursive Feature Elimination Method" adalah gagasan asli dari saya sendiri dan belum pernah dijadikan Tugas Akhir atau sejenisnya di Universitas Islam Negeri Sultan Syarif Kasim Riau maupun di perguruan tinggi lain.
2. Dalam Tugas Akhir ini TIDAK terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali tertulis dengan jelas dan dicantumkan sebagai referensi di dalam Daftar Pustaka.
3. Dalam Tugas Akhir ini TIDAK terdapat penggunaan Kecerdasan Buatan Generatif (Generative AI) yang bertentangan dengan ketentuan dan peraturan yang berlaku.
4. Saya bersedia menerima sanksi sesuai ketentuan yang berlaku apabila di kemudian hari terbukti bahwa Tugas Akhir ini melanggar kode etik maupun peraturan yang berlaku, termasuk plagiat ataupun pelanggaran hak cipta.

Demikianlah pernyataan ini dibuat dengan sebenarnya.

Pekanbaru, 12 Januari 2025

Yang membuat pernyataan,

CITRA WULANDARI

NIM.12150121717

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

Tugas akhir ini penulis persembahkan sebagai bentuk semangat, usaha, serta ungkapan cinta dan kasih sayang kepada orang-orang terpenting dalam hidup penulis. Dengan ketulusan hati dan rasa terima kasih yang mendalam, tugas akhir ini penulis persembahkan kepada:

1. Kedua orang tua tercinta, Bapak Suhendri dan Ibu Leni Hartati, serta adik-adik ku tersayang, juga seluruh keluarga besar penulis yang telah memberikan dukungan moral, materil, serta doa dan restu, sehingga penulis dapat menempuh pendidikan hingga jenjang S1 di Jurusan Teknik Informatika, UIN Sultan Syarif Kasim Riau.
2. Dosen pembimbing, Ibu Iis Afrianty, S.T., M.Sc., yang telah memberikan bimbingan, arahan, dan motivasi hingga tugas akhir ini dapat terselesaikan dengan baik
3. Seluruh dosen pengajar yang telah membimbing dan mendidik penulis dengan penuh kesabaran dan keikhlasan, sehingga ilmu yang diperoleh selama masa perkuliahan dapat menjadi bekal yang bermanfaat di masa depan.
4. Teman-teman seperjuangan di Program Studi Teknik Informatika, UIN Sultan Syarif Kasim Riau, atas kebersamaan dan dukungan selama menempuh perjalanan akademik.

Semoga tugas akhir ini dapat memberikan manfaat bagi para pembaca. Aamiin ya Rabbal ‘Alamiin.

Bit-Tech (Binary Digital - Technology)
Vol.8, No.2, December 2025
Available online at: http://jurnal.kdi.or.id/index.php/bt

Thyroid Disease Classification Using Support Vector Machine and Recursive Feature Elimination Method

Citra Wulandari[1)], Iis Afrianty[2)*], Elvia Budianita[3)], Siska Kurnia Gusti[4)]

[1)2)3)4)] *Informatic, Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia*

[1)] 12150121717@studens.uin-suska.ac.id

[2)] iis.afrianty@uin-suska.ac.id

[3)] elvia.budianita@uin-suska.ac.id

[4)] *siskakurniagusti@uin-suska.ac.id*

Article history:

Received 05 Dec 2025;
Revised 09 Dec 2025;
Accepted 10 Dec 2025;
Available online 10 Dec 2025

Keywords:

ADASYN
Machine Learning
Recursive Feature Elimination
Support Vector Machine
Thyroid Disease Classfication

Abstract

Thyroid disease is a common endocrine disorder that can cause serious metabolic and cardiovascular complications, so accurate early detection is clinically essential. This study proposes a Support Vector Machine (SVM) classifier enhanced with Recursive Feature Elimination (RFE) to select the most informative attributes and Adaptive Synthetic Sampling (ADASYN) to handle class imbalance in a Kaggle thyroid dataset of 3,771 clinical records. The data contain 25 diagnostic attributes with a strongly skewed distribution between healthy and thyroid cases. The model's robustness was examined using three train–test split ratios. The best configuration, SVM with a Linear kernel and 20 RFE-selected features under an 80:20 split, achieved 98.39% accuracy, with precision, recall, and F1-score all reaching 0.98, indicating consistently strong performance across classes. RFE contributes by removing redundant or weakly relevant variables, helping the classifier construct a more stable and interpretable decision boundary. ADASYN further improves the representation of the minority class, yielding higher recall and F1-score for thyroid cases and reducing the risk of missed diagnoses. Overall, the combined use of feature selection and adaptive oversampling produces a balanced and computationally efficient model for thyroid disease classification. These findings suggest that the proposed approach can support clinical decision-making, reduce diagnostic errors in imbalanced data settings, and strengthen early detection efforts in endocrine health assessment. By offering high sensitivity for thyroid cases while maintaining robust specificity for healthy patients, the model is well suited for integration into clinical decision-support and routine screening workflows.

I. INTRODUCTION

Thyroid disease is one of the most common hormonal disorders worldwide, affecting millions of people across all age groups and genders. Globally, approximately 5% of the population suffers from thyroid disorders, and in women, this figure can reach up to 10% due to hormonal fluctuations that influence thyroid function [1]. In clinical practice, this substantial and frequently underdiagnosed burden highlights the need for diagnostic approaches that can support earlier and more consistent detection of thyroid dysfunction, particularly in primary care and routine screening settings. In Indonesia, thyroid nodules are quite prevalent, and their detection has increased in recent years due to better public awareness and improved diagnostic imaging technologies. Recent hospital-based observations also suggest a steady rise in documented thyroid cases over time, especially in urban populations, indicating an increasing endocrine disease burden at the national level. Studies indicate that individuals aged 40–49 years are the most frequently affected group, representing roughly 32% of total cases, suggesting that middle-aged adults are particularly vulnerable to thyroid dysfunction [2]. Thyroid disorders not only impact hormone production but also have systemic effects on the body. These systemic disturbances can lead to functional impairment and reduced quality of life, further reinforcing the clinical importance of timely diagnosis and risk stratification. These conditions can increase the risk of serious cardiovascular complications such as coronary

* Corresponding author

ISSN 2622-2728 (online) 2622-271X (print)
doi: 10.32877/bt.v8i2.3454

heart disease, heart failure, and stroke, emphasizing the importance of early diagnosis and proper management [3]. The thyroid gland produces several essential hormones, namely Thyroxine (T4), Triiodothyronine (T3), and Thyroid Stimulating Hormone (TSH), which collectively regulate metabolism, energy production, and overall physiological functions. Imbalances in these hormones are key indicators of thyroid disorders and are critical for accurate diagnosis [4]. To date, no treatment exists that can completely cure thyroid diseases. Current therapies primarily aim to manage symptoms and stabilize hormone levels. For example, hypothyroidism is often treated with levothyroxine, while hyperthyroidism may be managed with antithyroid drugs, radioiodine therapy, or thyroidectomy,Although these treatments help regulate hormone levels, they do not repair or restore damaged thyroid tissue [5].The complexity of thyroid disorders arises from multiple causative factors, including autoimmune reactions, genetic predisposition, and iodine deficiency, alongside limited availability of long-term follow-up data, which complicates the determination of optimal individualized treatment strategies [6]. Recently, machine learning technologies have been increasingly applied in healthcare, particularly for early detection and classification of thyroid disorders. Machine learning algorithms can analyze complex patterns in clinical and laboratory data more efficiently and accurately than traditional diagnostic methods [7]. Within this context, the present study explores a classification framework based on Support Vector Machine (SVM) combined with Recursive Feature Elimination (RFE) and Adaptive Synthetic Sampling (ADASYN), with the explicit aim of developing a data-driven diagnostic tool that can be aligned with clinical workflows for thyroid disease screening and decision support.

Robust scaler is a data preprocessing technique used to normalize features in a way that is more resistant to outlier values [8]. Unlike scaling techniques such as standard scaler or min-max scaler, which use mean, standard deviation, minimum, or maximum values that are easily affected by outliers. Instead, robust scaler works by reducing data values to the median and dividing them by the interquartile range (IQR). By using Q1, Q2 (median), and Q3, this technique prevents data from changing drastically due to deviating values. Because it only utilizes the middle part of the data, robust scaler is the right choice for datasets with many outliers, so that the scaling results remain stable and are not easily disturbed. In the context of thyroid disease data, where laboratory indicators such as TSH, TT4, or FTI can exhibit extreme values, robust scaling helps stabilize the feature space so that the subsequent SVM classifier can construct a more reliable decision boundary.

Adaptive Synthetic Sampling (ADASYN) is an oversampling method developed to address class imbalance by adaptively generating synthetic samples in the most difficult-to-learn minority class areas. Unlike conventional oversampling methods, ADASYN places greater emphasis on minority samples with high classification error rates, enabling the model to learn minority patterns more effectively. Research shows that the application of ADASYN in stroke disease classification successfully improves recall and F1-score values because the class distribution becomes more balanced [9]. By focusing on hard-to-classify minority instances, ADASYN is particularly relevant for thyroid datasets where confirmed disease cases are much fewer than healthy cases, as it can enhance sensitivity to thyroid disorders and reduce the likelihood of missed diagnoses in imbalanced clinical settings.

Among these algorithms, Support Vector Machine (SVM) has gained prominence as a supervised learning technique capable of both classification and regression tasks. SVM aims to identify the optimal hyperplane that separates data from different classes with the maximum margin, making it highly effective for high-dimensional datasets [10]. SVM can handle both linear and non-linear data through the use of kernel functions, which map input features into higher-dimensional spaces to enhance separability [11]. However, SVM performance strongly depends on the quality and relevance of input data. Datasets that are unbalanced or contain irrelevant features may reduce the model's accuracy and stability, necessitating the use of supporting techniques to improve classification results [12]. Therefore, supporting methods are needed to improve SVM performance and classification results.

Two commonly employed techniques to enhance SVM performance are Recursive Feature Elimination (RFE) and Adaptive Synthetic Sampling (ADASYN). RFE is a feature selection method that iteratively removes less important features based on their contribution to the model, retaining only the most relevant subset. This reduces model complexity while maintaining or improving predictive accuracy [13]. ADASYN addresses the challenge of imbalanced datasets by adaptively generating synthetic samples for minority classes, enabling the model to better learn rare patterns and improve classification performance in difficult to classify cases [14]. The combination of SVM with RFE and ADASYN has demonstrated improved stability, robustness, and accuracy in various medical classification tasks, including stroke prediction, cancer detection, and thyroid disease classification [15]. By selecting the most informative features and balancing class distributions, the model can achieve higher precision, recall, and F1-score, which are critical for evaluating the effectiveness of diagnostic tools.

Based on this background, this study aims to investigate the application of an integrated approach combining SVM, RFE, and ADASYN for thyroid disease classification. Specifically, it seeks to address the identified research gap by providing a systematic evaluation of SVM configurations with and without RFE and ADASYN on an imbalanced thyroid dataset. The research focuses on evaluating how RFE can enhance SVM effectiveness through optimal feature selection and how ADASYN can mitigate data imbalance issues. In particular, the study

quantifies the impact of each technique individually and in combination on key performance indicators such as accuracy, precision, recall, and F1-score, with special attention to improvements in minority-class detection. By leveraging these methods, the study seeks to improve model performance in terms of accuracy, precision, recall, and F1-score. From a practical standpoint, the proposed framework is intended not only as a technical contribution but also as a candidate decision-support component that can be integrated into clinical information systems to flag patients at higher risk of thyroid dysfunction. This approach is expected to provide a more reliable and interpretable classification model that can assist healthcare professionals in early detection, improve treatment planning, and ultimately reduce the risk of severe complications associated with thyroid disorders.

II. Related Works/Literature Review

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression. This algorithm works by finding the best separating boundary so that data from each class can be clearly distinguished [10]. Research by [16] Performing a comparison of methods for predicting thyroid disease with a dataset of 3,772 cases and 22 variables. Several classification methods were compared, including Support Vector Machine (SVM), Random Forest, Decision Tree, Logistic Regression (LR), and K-Nearest Neighbors (KNN). The SVM method had the highest accuracy of 98.6%, with a precision of 98.4%, recall of 98.7%, and an F1-score of 98.5%.

Similar research conducted by [17] which compares several methods for predicting thyroid disease with 422 data points and 26 variables. The research dataset is unbalanced, with 125 cases of thyroid disease and 297 cases of no thyroid disease. This study compares methods such as Random Forest, Decision Tree, and Support Vector Machine. The SVM method has the highest accuracy of 90%. These findings confirm that SVM is a strong baseline for thyroid disease classification, but they also show that performance can be affected by dataset characteristics such as imbalance and noise. Previous studies have shown that the SVM method can deliver fairly good results, but the outcomes may vary depending on the characteristics of the dataset and the testing method used. Some studies also achieved less-than-optimal accuracy due to the imbalance in the number of data points between classes. Therefore, this study will incorporate additional methods to address this issue, aiming to make thyroid disease predictions more accurate and consistent.

Robust Scaler is one of the important preprocessing techniques for addressing scale differences and sensitivity to outliers. Research by [18]. The choice of scaling technique greatly affects the performance of classification models, especially in algorithms and datasets with varying levels of imbalance. Of the five scaling techniques tested, the robust scaler showed superiority because it is resistant to outliers and effective for data with uneven distribution. The study also emphasized that each algorithm has a different level of sensitivity to changes in data scale, so choosing a scaling technique such as the robust scaler is the right step in the machine learning process. In thyroid disease datasets, which often contain extreme laboratory values, robust scaling helps stabilize feature ranges so that downstream models such as SVM can build more reliable separating hyperplanes.

Feature selection (RFE) is a method that works gradually by discarding features with small contributions until only the most important features remain. This method has been proven to make models more efficient without reducing prediction accuracy [19]. Research conducted by [13] Using a dataset consisting of 569 samples and 30 variables, the results of the study show the use of the RFE method in SVM for breast cancer classification. After feature selection using RFE, the number of features was successfully reduced to the 15 most important features without reducing the model's performance. From the test results, the best model with a 90:10 data split produced an accuracy value of 98%, precision of 100%, recall of 94%, and an F1-score of 97% [13]. Therefore, RFE is an important component in improving the efficiency and stability of medical classification models. These RFE-based studies illustrate the strand of literature that focuses on feature selection as a way to reduce dimensionality, remove noisy attributes, and preserve or even enhance diagnostic performance in clinical prediction tasks. In addition to feature selection, data balancing is also very important for improving SVM performance. ADASYN is an adaptation of the SMOTE method that works in a more adaptive way. This technique creates new synthetic data around the most difficult-to-predict minority samples, allowing the model to learn better and be balanced across all classes [14]. This method can improve data distribution and increase the sensitivity of the model to minority classes without causing excessive overfitting [20] [21]. A similar study conducted by [15] Applying SMOTE to stroke disease classification using SVM with RBF kernel. This combination improves accuracy to 90.51% and significantly increases precision and recall values. Because ADASYN works more adaptively than SMOTE, this approach is considered more effective for medical datasets with unbalanced distributions, such as thyroid disease cases. Taken together, these works represent the resampling and data-balancing line of research, showing that oversampling techniques can substantially improve minority-class detection when used alongside SVM in medical settings.

Several previous studies have also compared SVM with other algorithms in the context of medical classification. Research by [12] shows that SVM with RBF kernel provides better results than logistic regression in predicting stroke, with an accuracy of 84%. Research by [22] Using SVM to diagnose chronic kidney disease and achieving an accuracy of 96.42%. However, the two previous studies did not use resampling and feature

selection techniques. The latest study shows that combining RFE and ADASYN can make the model stronger and improve predictive capabilities in complex medical data [23]. These findings indicate that SVM-based models, feature selection techniques such as RFE, and resampling methods such as ADASYN or SMOTE have each been studied, but most existing work treats them in isolation or in partial combinations rather than as a fully integrated framework evaluated on thyroid disease data.

Based on the discussion, it is known that SVM performance is greatly influenced by feature relevance and data balance. RFE and ADASYN provide complementary benefits, as RFE filters out features that are truly needed, while ADASYN helps to even out the distribution of data for each class. However, the simultaneous and systematic use of these two methods in a single SVM-based framework for thyroid disease classification especially on large, highly imbalanced clinical datasets has been rarely explored in previous studies. Therefore, this study aims to build an SVM model using RFE and ADASYN. The novelty of this research lies in explicitly evaluating the individual and combined effects of RFE and ADASYN on SVM performance for thyroid disease, with a focus on key indicators such as precision, recall, and F1-score for the minority (diseased) class. This study produces a more accurate and consistent disease classification system that is capable of handling unbalanced data. In addition, this study is expected to contribute to the development of machine learning-based medical diagnosis technology in Indonesia.

III. METHODS

This quantitative study evaluates the impact of feature selection and data balancing on the performance of SVM in classifying thyroid disease. The workflow includes data collection, preprocessing, transformation, imbalanced data handling, RFE application, data splitting, and SVM training and testing. These steps were conducted to obtain an optimal classification model for medical data, as illustrated in Figure 1.

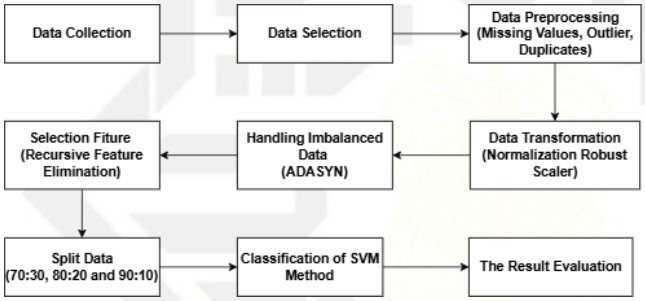


Fig. 1 Research Methodology

Figure 1 presents the research flow consisting of data selection, preprocessing, transformation, class balancing with ADASYN, feature selection using RFE, and SVM modeling. The final stage evaluates model performance using accuracy, precision, recall, and F1-score.

A. Data Collection

The data used in this study is secondary data in the form of a collection of datasets obtained from the Kaggle platform. The data can be accessed via https://www.kaggle.com/datasets/yasserhessein/thyroid-disease-data-set. The Thyroid Disease Data Set consists of 3,771 data points with 25 attributes.

TABLE 1
ORIGINAL DATASET

No	Age	Sex	...	TSH	TT4	T4U	FTI	BINARY CLASS
1	41	0	...	1.3	125	1.14	109	0
2	23	0	...	4.1	102	0.995	110.4696	0
3	46	1	...	0.98	109	0.91	120	0
4	70	0	...	0.16	175	0.995	110.4696	0
5	70	0	...	0.72	61	0.87	70	0
...	...	...	...	...	...	...	...	...
3768	68	0	...	1	124	1.08	114	0
3769	74	0	...	5.1	112	1.07	105	0
3770	72	1	...	0.7	82	0.94	87	0
3771	64	0	...	1	99	1.07	92	0

Table 1 shows the original data values containing patient information, such as age, sex, and thyroid hormone test values. The binary class column is used as a label to distinguish between patients who are indicated to have thyroid disease and those who are not. This data forms the basis for further analysis and modeling.

TABLE 2
DATASET ATTRIBUTES

No	Attribute	Description
1	Age	Patient's Age
2	Sex	Patient's Gender
3	On thyroxine	Currently Taking Thyroxine Medication
4	Query on thyroxine	History of Thyroxine Usage
5	On_antithyroid medication	Currently Taking Anti-Thyroid Medication
6	Sick	Presence of Illness
7	Pregnant	Pregnancy Status
8	Thyroid surgery	History of Thyroid Surgery
9	1131 treatment	History of Radioactive Iodine (I-131) Treatment
10	Query hypothyroid	Suspected Hypothyroidism
11	Query hyperthyroid	Suspected Hyperthyroidism
12	Lithium	Lithium Consumption
13	Goiter	Presence of Goiter or Thyroid Enlargement
14	Tumor	Presence of Tumor
15	Hypopituitary	Pituitary Gland Disorder
16	Psych	Psychological Disorder
17	TSH measured	TSH Measurement Status
18	TSH	Thyroid-Stimulating Hormone Level
19	T3 measured	T3 Measurement Status
20	T3	Triiodothyronine Level
21	TT4 measured	Total Thyroxine (TT4) Measurement Status
22	TT4	Total Thyroxine Level
23	T4U measured	T4 Uptake Measurement Status
24	T4U	T4 Uptake Value
25	FTI	Free Thyroxine Index Value

Table 2 shows all attributes used in the thyroid disease dataset along with their descriptions. These attributes form the basis for the preprocessing, feature selection, and classification model development processes in this study.

B. Data Selection

In the data selection stage, the research begins by identifying and retaining only the most relevant attributes and samples to ensure the integrity and representativeness of the dataset used for classification. This step is essential for minimizing noise and reducing unnecessary model complexity, particularly in medical datasets where not all variables contribute equally to diagnostic decision-making. The dataset employed in this study consists of 3,771 samples and 25 clinical and laboratory attributes related to thyroid function. Each attribute is examined for its clinical relevance and suitability for inclusion in the modeling workflow to avoid incorporating irrelevant or misleading information.

Furthermore, data selection ensures that the final dataset accurately reflects the underlying population characteristics, enabling the classification model to generalize effectively. By validating the completeness of attributes and the consistency of samples across all entries, the process reduces the risk of bias arising from non-representative variables or inconsistent data patterns. This stage also prepares the dataset for subsequent pre-processing and transformation steps, creating a structured and coherent foundation for building a reliable and accurate classification model.

C. Data Pre-processing

The pre-processing stage is carried out to ensure that the dataset is clean, coherent, and suitable for machine learning model development. One of the primary tasks in this stage involves handling missing values that, if ignored, may introduce significant bias or distort the model's learning process. Depending on the extent and pattern of missingness, appropriate strategies such as imputation or removal of problematic rows are applied. Additionally, outlier detection is conducted to identify extreme values that could disrupt distributional assumptions or disproportionately influence the classification boundary.

Another crucial aspect of pre-processing is the removal of duplicated records and the reorganization of the dataset to ensure structural consistency. Duplicate entries can distort the model's perception of class distributions, potentially leading to overfitting and reduced generalization ability. By ensuring that each record is unique and that the dataset maintains a coherent and logical structure, this stage enhances the reliability of the learning process. Ultimately, effective pre-processing improves data quality and provides a stable foundation for the subsequent modeling phases, contributing directly to better predictive performance..

D. Data Transformation

Data transformation is conducted to standardize the scale of all variables so that the classification model, particularly SVM, can operate more effectively. In this study, the Robust Scaler method is employed due to its resilience against outliers, which are commonly present in clinical laboratory datasets. This technique normalizes

feature values by subtracting the median (Q2) and dividing by the interquartile range (Q3–Q1), as represented in Equation (1). Unlike standard normalization approaches that rely on the mean and standard deviation, Robust Scaler focuses on the central portion of the data, thereby preventing extreme values from disproportionately affecting the scaling process.

$$X' = \frac{(X - Median\,(X))}{(Q3 - Q1)} \quad (1)$$

Where Q2 is the median, and Q1 and Q3 are the first and third quartiles, respectively. This method has been proven effective in maintaining data scale stability and improving classification model accuracy in datasets containing outliers [24].

E. Handling Imbalanced Data

ADASYN is a technique used to add data to classes with small numbers so that the dataset becomes more balanced Unlike SMOTE, which adds data evenly, ADASYN creates new data mainly in minority classes that are difficult to learn. Figure 2 shows the data before ADASYN. There are a total of 3,417 data points in class 0 (no thyroid) and 291 data points in class 1 (thyroid) out of a total of 3,771 data points.

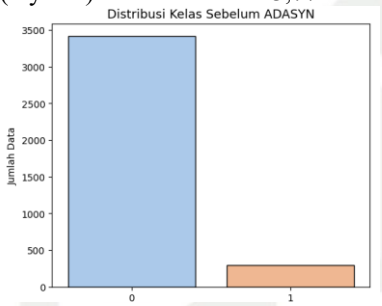


Fig. 2 Data Before ADASYN

Figure 3 shows the data after applying Adasyn. The data is balanced with a total of 6,817 samples, consisting of 3,417 for class 0 and 3,400 for class 1. With this balance, the model can detect patterns in both classes more effectively.

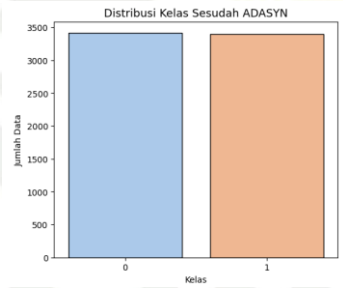


Fig. 3 Data After Adasyn

ADASYN works by adaptively adding synthetic data to minority classes, especially to the most difficult samples to learn. The difficulty level is calculated from the number of majority neighbors around the minority sample, then new data is created through interpolation with the nearest neighbors. This process helps balance the class distribution so that the model can better recognize minority patterns [25].

F. Selection Feature (RFE)

Recursive Feature Elimination (RFE) is a wrapper-based feature selection method that works by repeatedly training the model, then gradually eliminating features with the lowest contribution until a subset of the most relevant features is obtained to improve prediction performance[19]. In this study, RFE was applied to minimize the number of irrelevant or redundant features, so that the Support Vector Machine (SVM) classification model could run more efficiently, faster, and achieve a better level of accuracy. The formula for determining the weights of these features is defined by [26].

At this stage, the feature weights are computed using the separating hyperplane of the linear SVM. Each weight is derived from the contribution of individual training samples through the Lagrange multiplier (α_i), class label (y_i), and feature vector (x_i), as expressed in Equation (2):

$$w = \sum_{k=1}^{n} \alpha_k y_k x_k \quad (2)$$

Subsequently, the ranking criterion for each feature is computed by squaring the corresponding weight component. This criterion determines the order in which features are eliminated, with features having the smallest weight magnitude removed first. The ranking calculation is presented in Equation (3):

$$c_k = (w_k)^2,\ k = 1, 2, \dots, |S| \quad (3)$$

In the third stage, features are sorted based on their weight, where features with the lowest weight are eliminated in each iteration. The Recursive Feature Elimination (RFE) process involves retraining the linear SVM

model in each iteration. Next, the SVM model is retrained using the remaining features, and this procedure is repeated until all features are eliminated. At the end of the process, the features are sorted based on the order of elimination, with the last eliminated feature considered to have the most significant influence [27].

Weight w is the SVM separating vector the value a_i is the Lagrange multiplier for the i-th data point. Wher y_i is the class label. The feature vector of the-ith data point is denoted as x_i. Feature index marked as k, and the total number of features in the dataset is denoted as $|S|$. The value c_k is the ranking criterion value for feature k, and $(w_k)^2$ is the weight contribution of that feature.

G. Data Splitting

The data splitting stage is a critical component of the machine learning workflow, as it separates the dataset into training and testing subsets to enable systematic model development and evaluation. The training set is used to fit the model and learn the decision boundaries, while the testing set serves as independent data to assess how well the model generalizes to previously unseen cases. This separation is essential for preventing information leakage, which can occur when the model inadvertently learns patterns from the test set, leading to inflated performance metrics. By ensuring that the model is evaluated solely on data it has not encountered during training, the data splitting process provides a more reliable estimate of the model's real-world classification capability.

In this study, the dataset is divided using common ratios such as 90:10, 80:20, and 70:30 to examine the effects of different training sizes on model stability and predictive performance. Larger training portions, such as the 90:10 split, provide the model with more data to learn complex relationships, whereas splits like 70:30 offer a larger test set for more rigorous evaluation. Testing across multiple ratios allows for a comprehensive analysis of model behavior under varying conditions, helping determine the most optimal balance between training depth and generalization strength. This approach ensures a more robust evaluation framework and reduces the risk of overfitting, ultimately supporting the development of a classifier that performs consistently across diverse data distributions.

H. SVM Classification Method

At this stage, the SVM model is trained and tested using thyroid data that has been split into 90:10, 80:20, and 70:30 ratios. Classification is performed with linear, polynomial, and RBF kernels, each tested with variations of the C parameter (1, 10, 100). The polynomial kernel uses degrees 1, 2, and 3, while the RBF kernel is evaluated with gamma settings (scale and gamma) and gamma values of 1, 2, and 3. The linear kernel is defined in Equation (4) as a simple dot product between feature vectors.:

$$K\ (x_{i,}\ x_j\ ) = x_i .\ x_j \quad (4)$$

The polynomial kernel used in this study follows the formulation presented in Equation (5), where the kernel function is expressed as:

$$K\ (x_{i,}\ x_j\ ) = (\ x_i .\ x_j + c)^{\mathrm{d}} \quad (5)$$

In this equation, $d$represents the polynomial degree and $c$is a constant term controlling model flexibility. Meanwhile, the RBF kernel is computed according to Equation (6):

$$K(x_i, x_j) = \exp(\ -\gamma\ ||x_{i-}\ x_j||\ )^2 \quad (6)$$

In a polynomial kernel, the parameter d is the degree of the polynomial and c is the constant. The parameter γ used in the RBF kernel as a regulator of width of the gaussian function.

I. Confusion Matrix Evaluation

A confusion matrix is a useful tool for evaluating classification performance by comparing the model's predictions with the actual data. It provides a clear overview of how accurately the model recognizes each class [28]. A confusion matrix is a concept in machine learning that studies existing data and groups it into new data by generating output in the form of categorical variables, both nominal and ordinal [29]. A confusion matrix consists of four key components: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Accuracy is a measure used to see how well a model can classify with correct results. Accuracy can be calculated using Equation (7).

$$\text{Accuracy} = \frac{(TP+TN)}{TP+TN+FP+FN} \text{ x } 100\% \quad (7)$$

1) Precision is a measure that shows how accurate the model is in predicting positive data. Precision can be calculated using the equation (8).

$$\text{Precision} = \frac{(TP+TN)}{TP+FP} \text{ x } 100\% \quad (8)$$

2) Recall is a measure that shows how well the model recognizes actual positive data. Recall can be calculated using Equation (9)

$$\text{Recall} = \frac{TP}{TP+FN} \text{ x } 100\% \quad (9)$$

3) The F1 score is a measure used to assess the balance between precision and recall. The F1 score can be calculated using Equation (10).

$$\text{F1-Score} = \frac{2\ (Precision\ x\ Recall)}{(Precision+Recall)} \quad (10)$$

IV. RESULTS

This study used 3,771 cleaned thyroid records with 25 variables and was implemented in Python using Google Colab. RFE was applied for feature selection, while SVM with linear, RBF, and polynomial kernels served as the classification model. Performance was evaluated using the confusion matrix across three data split scenarios (70:30, 80:20, 90:10), with accuracy results presented in tabular form for comparison.

A. Data Prepocessing

In this study, the preprocessing stage was conducted through three primary steps: handling outlier values, managing missing values, and removing duplicate records. Prior to outlier correction, a descriptive statistical assessment was performed to understand the distribution of the main numerical attributes, namely Age, TSH, TT4, T4U, and FTI. The dataset consisted of 3,708 valid entries for each of these features. The Age variable exhibited a mean of 51.74 years with a standard deviation of 19.00, ranging from a minimum of 1 year to a maximum of 94 years. Similarly, the TSH values showed a wide dispersion, with a mean of 5.08 and a standard deviation of 23.49, spanning from 0.005 at the minimum to an extreme maximum of 530. TT4 also displayed substantial variability, with values ranging from 2 to 430 and a mean of 108.32. For T4U, the values ranged from 0.25 to 2.32 with a relatively narrow mean of 0.99, while FTI varied between 2 and 395 with a mean of 110.48.

The presence of large gaps between minimum and maximum values particularly in TT4, TSH, and FTI indicates clear outlier behavior that could distort model learning if left unaddressed. Quartile analysis further supports this observation, with the 25th, 50th, and 75th percentiles showing far more concentrated ranges compared to the extreme maximum values. For example, TT4 exhibited quartiles of 88.75, 105, and 123, while its maximum reached 430. Likewise, TSH had quartiles of 0.58, 1.55, and 3.50, yet an extreme maximum outlier of 530. These discrepancies highlight the need for thorough preprocessing to ensure data stability. Consequently, outlier handling procedures were implemented to reduce noise and improve the reliability of subsequent modeling. Cleaning these anomalies ensures that the classification model can learn more representative patterns from the thyroid dataset, ultimately supporting more accurate and robust prediction performance.

1) Data Transformation

In this step, Robust Scaler normalization is used. The normalized data can be seen in Table 3. The calculation process can be seen in Equation (1).

TABEL 3
DATA NORMALIZATION

No	Age	Sex	...	TSH	TT4	T4U	FTI
1	-0.419355	0	...	-0.085616	0.583942	0.833333	-0.03571
2	-1	0	...	0.873288	-0.087591	0.027776	0.01677
3	-0.258065	1	...	-0.195205	0.116788	-0.444444	0.35714
...	...	...	...	...	...	...	..
3769	0.645161	0	...	1.215753	0.20438	0.444444	-0.17857
3770	0.580645	1	...	-0.291096	-0.671533	-0.277778	-0.82142
3771	0.322581	0	...	-0.188356	-0.175182	0.444444	-0.64285

B. Recursive Feature Elimination (RFE)

The result of the feature selection process using RFE is a ranking of all attributes from the most influential to the least influential. In addition, RFE also produces a list of selected attributes according to the specified number of features. Table 4 shows the calculation results and feature rankings based on the RFE method. The calculations for the RFE method can be seen in Equation (2) and Equation (3).

TABLE 4
RFE SELECTION FEATURE CALCULATION RESULTS

No	FEATURE NAME	SELECTION STATUS	WEIGHT VALUE
1	On thyroxine	Selected	7.6928
2	Thyroid surgery	Selected	5
3	On antithyroid medication	Selected	2
4	I131 treatment	Selected	1.3609
5	TSH measured	Selected	1.2718

6	Psych	Selected	0.9401
7	Query hyperthyroid	Selected	0.8010
8	Query on thyroxine	Selected	0.7643
9	FTI measured	Selected	0.7533
10	T4U measured	Selected	0.7502
11	Lithium	Selected	0.7090
12	Query Hypothyroid	Selected	0.5892
13	Tumor	Selected	0.5603
14	Sick	Selected	0.5098
15	TSH	Selected	0.4706
16	Goitre	Selected	0.4276
17	T3 measured	Selected	0.1922
18	Pregnant	Selected	0.1501
19	T4U	Selected	0.0094
20	Sex	Selected	0.0038
21	Age	Eliminated	0
22	Hypopituitary	Eliminated	0
23	TT4 measured	Eliminated	0
24	TT4	Eliminated	0
25	FTI	Eliminated	0

The RFE results across all scenarios show that several features such as Pregnant, Age, Hypopituitary, TT4 measured, TT4, and FTI were eliminated because they had low weights in the formation of the decision boundary line in SVM. This pattern is consistent with how RFE works, which repeatedly trains the model, assesses the weight of each feature, and then removes the features with the smallest contribution. As a result, only important features are retained, making the SVM model more efficient, more resistant to noise, and producing more stable accuracy across various data partitioning schemes.

C. Modeling With SVM

SVM was applied for classification across four scenarios to examine the effects of balancing and featureselection using linear, polynomial, and RBF kernels. The test scenarios are listed in Table 5.

TABLE 5
TEST SCENARIO

Data Sharing	Method	Features
70:30, 80:20, 90:10	SVM	25
70:30, 80:20, 90:10	SVM + ADASYN	25
70:30, 80:20, 90:10	SVM + RFE	20, 15, and 10
70:30, 80:20, 90:10	SVM + RFE + ADASYN	20, 15, and 10

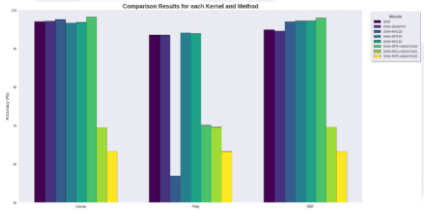


Fig. 4 Comparison Results for Each Kernel

The Figure 4 compares the accuracy of three SVM kernels across multiple methods, SVM, SVM with ADASYN, SVM with RFE, and their combination. Linear and RBF kernels outperform the Polynomial kernel. Performance may increase or decrease with ADASYN and RFE, but the best result is achieved by SVM + RFE + ADASYN using 20 features with the Linear kernel, reaching 98.19% accuracy.During the evaluation stage, model performance was compared based on four groups of methods, namely SVM, SVM + ADASYN, SVM + RFE, and SVM,RFE, and ADASYN. The best accuracy results for the SVM method without balancing and without feature selection were obtained in the 80:20 data split using a linear kernel, with an accuracy of 97.17%. The highest accuracy was achieved by the SVM and RFE method at an 80:20 ratio and also using a linear kernel, with an accuracy value of 98.39%.

The confusion matrix results show that feature reduction through RFE can improve model stability and precision by significantly reducing the number of negative negatives. The last method, SVM, RFE, and ADASYN, achieved the highest accuracy at a 70:30 ratio and also used a linear kernel, with an accuracy value of 98.17%. This shows that the combination of data balancing and feature selection produces the most optimal performance improvement. Overall, the evaluation results show that the linear kernel consistently provides the best performance across all method groups, and the use of RFE and ADASYN makes the model more accurate and balanced, especially in terms of minority class recognition.

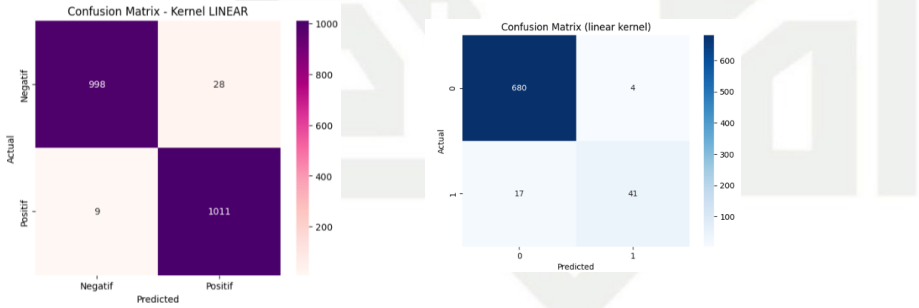
D. Evaluation

After modeling, the SVM model was trained and tested using four scenarios evaluated through accuracy, precision, recall, and F1-score. These metrics, calculated using the confusion matrix, measure the model's ability to distinguish thyroid from non-thyroid cases. The results for the 70:30, 80:20, and 90:10 data splits are presented in Table 6.

TABLE 6
PERFORMANCE COMPARISON AND SELECTION OF HIGHEST ACCURACY IN VARIOUS CONFIGURATIONS

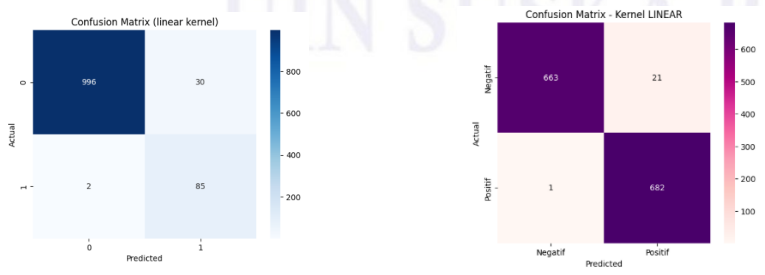
Method	Feature	Kernel	Accuracy	Precisiom	Recall	F1- Score
SVM	25	Linear	97.17%	97%	97%	97%
		Polynomial	93.67%	93%	94%	92%
		RBF	94.47%	94%	94%	93%
SVM + ADASYN	25	Linear	96.77%	98%	97%	97%
		Polynomial	92.72%	96%	93%	94%
		RBF	94.07%	96%	94%	95%
SVM + RFE	20	Linear	98.39%	98%	98%	98%
		Polynomial	62.62%	78%	63%	57%
		RBF	97.51%	98%	98%	98%
SVM + RFE	15	Linear	96.68%	97%	97%	97%
		Polynomial	94.07%	94%	94%	93%
		RBF	97.21%	97%	97%	97%
SVM + RFE	10	Linear	96.86%	97%	97%	97%
		Polynomial	93.89%	94%	94%	92%
		RBF	97.21%	97%	97%	97%
SVM + RFE + ADASYN	20	Linear	98.19%	98%	98%	98%
		Polynomial	70.04%	80%	70%	67%
		RBF	97.95%	98%	98%	98%
SVM + RFE + ADASYN	15	Linear	97.07%	97%	97%	97%
		Polynomial	71.41%	81%	71%	69%
		RBF	97.65%	98%	98%	98%
SVM + RFE + ADASYN	10	Linear	63.71%	73%	64%	60%
		Polynomial	63.71%	73%	64%	60%
		RBF	63.64%	73%	64%	60%

SVM + RFE with 20 features produced the highest accuracy, namely 98.39% on the Linear kernel. RFE proved to be effective in selecting the most influential features so that the model was more efficient and accurate. SVM without RFE or with ADASYN remains good, but does not exceed SVM + RFE because it still uses all the initial features. ADASYN is used for evaluation on balanced data and the results confirm that optimal feature selection gives the best performance. A comparison of the accuracy of each kernel and its calculation is shown in Equations (4), (5), and (6).



(a) (b)

Fig. 5 Comparison of confusion matrices with the highest accuracy in SVM testing
(a) SVM without RFE or ADASYN, and (b) Combination Method (SVM + RFE + ADASYN)



(c) (d)

Fig. 6 Comparison of confusion matrix result with the highest accuracy in SVM testing

(c)SVM+ ADASYN, dan (d) SVM +RFE

The comparison of confusion matrices presented in Figure 5 and Figure 6 provides a detailed visualization of how different SVM configurations perform under various preprocessing and feature selection strategies. Figure 5(a) illustrates the confusion matrix of the baseline SVM model without RFE or ADASYN, showing that although overall accuracy is high, the model still misclassifies several minority-class samples due to the imbalance in the original dataset. In contrast, Figure 5(b), which corresponds to the combined SVM + RFE + ADASYN method, demonstrates a marked improvement in correctly identifying positive thyroid cases, indicating that integrating feature reduction and adaptive oversampling significantly enhances minority-class sensitivity. Meanwhile, Figure 6(c) displays the results for SVM + ADASYN, showing improved recall but slightly reduced specificity due to the synthetic oversampling process, whereas Figure 6(d), representing SVM + RFE, highlights the stabilizing effect of feature elimination in reducing false negatives while maintaining high precision. Overall, the comparison across Figures 5 and 6 confirms that the best-performing configuration is the integration of RFE and ADASYN, as it produces a more balanced classification outcome and minimizes misclassification in both majority and minority classes.

V. DISCUSSION

The results indicate that the SVM + RFE configuration using a Linear kernel with an 80:20 data split achieved the highest accuracy of 98.39%. This outcome can be attributed to several interrelated technical factors involving the intrinsic characteristics of the selected features, the structural properties of the SVM algorithm, and the contribution of dimensionality reduction through RFE. Collectively, these elements interact to enhance the model's ability to construct a more discriminative and stable decision boundary, leading to improved overall classification performance. At the same time, these results should be interpreted as one promising but not definitive configuration, since different data sources or clinical settings may require recalibration of the model and reassessment of its assumptions.

First, RFE successfully filtered out the most relevant features, particularly clinical features such as TSH, T3, TT4, T4U, FTI, Age, and Sex, which contributed significantly to the SVM decision-making process. By eliminating features with low weight, the model became simpler and free from noise. Previous research by [13] also shows that RFE improves the stability and accuracy of classification models, especially in medical problems that have many features but not all of them are informative. Therefore, the increase in accuracy in this scenario supports the theory that removing irrelevant features can widen the hyperplane margin in SVM, thereby improving model performance. However, the present study did not compare RFE with alternative feature-selection strategies such as LASSO, embedded methods based on tree ensembles (e.g., Random Forest feature importance), or mutual-information-based filters, so it remains possible that other techniques could yield comparable or even superior feature subsets. Future work should therefore include a systematic comparison of multiple feature-selection approaches to determine whether the observed performance gains are specific to RFE or reflect a more general benefit of careful dimensionality reduction in thyroid disease classification.

Second, the Linear kernel proved to be the most stable in producing high accuracy because the relationship pattern between features in the thyroid dataset became more linear after undergoing Robust Scaler normalization. This study is in line with research [12] which shows that the Linear kernel is effective for medical data whose variables are measured directly and do not contain many non-linear patterns. Because Linear SVM forms a hyperplane without feature space transformation, the model becomes more stable against data variation, especially with a fairly balanced data ratio such as 80:20. Nonetheless, the preference for a Linear kernel in this work should not be generalized uncritically to all medical datasets, since scenarios with more complex non-linear relationships or multimodal inputs may favor non-linear kernels or alternative models such as gradient-boosted trees and deep neural networks.

Third, while ADASYN was applied, it did not provide notable gains, especially after RFE feature filtering. This is because the dataset's minority class was already well represented after preprocessing, so RFE became the primary factor driving accuracy improvements. Base on a comparison of method in the study [30]. ADASYN does not always produce the best result on all dataset. This show that the effectiveness of ADASYN is greatly influenced by the level of imbalance and characteristics of the dataset. Therefore, the high accuracy of SVM and RFE is not because ADASYN does not work, but because mathematically reducing relevant features has a greater impact on the stability of the SVM margin than adding synthetic samples. From a critical perspective, the limited incremental benefit of ADASYN in this study suggests that its impact may be more pronounced in settings with more extreme imbalance or higher noise levels than in the present dataset. Moreover, because ADASYN introduces synthetic minority instances in feature space, there is a non-trivial risk of overfitting to oversampled regions and of blurring the distinction between real and synthetic clinical patterns, which can complicate interpretability and clinicians' trust in the model's decisions. These considerations highlight the need for external validation on other imbalanced medical datasets and for careful monitoring of performance degradation when synthetic oversampling is applied beyond the original training distribution.

Fourth, the 80:20 data ratio provides a good balance between training data and test data. With a large amount of training data, SVM can form a more stable hyperplane, while the size of the test data is still sufficient to produce accurate evaluations. Because it is considered capable of balancing model complexity and generalization ability, the 80:20 ratio is often used in machine learning research. As a result, this ratio is one of the best configurations of the models tested in this study and provides consistent results. Even so, the evaluation protocol in this work is still limited to hold-out splits; additional experiments using k-fold cross-validation and independent external datasets would be valuable to further assess the robustness and generalizability of the proposed configuration.

VI. CONCLUSIONS

This study demonstrates that integrating Support Vector Machine (SVM) with Recursive Feature Elimination (RFE) substantially enhances the accuracy and stability of thyroid disease classification. The best performance, reaching an accuracy of 98.39%, was achieved when the SVM model utilized the 20 most influential features, including On thyroxine, thyroid surgery, on antithyroid medication, I131 treatment, TSH measured, Psych, Query hyperthyroid, Query on thyroxine, FTI measured, T4U measured, Lithium, Query hypothyroid, Tumor, Sick, TSH, Goitre, T3 measured, Pregnant, T4U, and Sex. The retention of these clinically relevant variables enables the classifier to construct a clearer and more discriminative hyperplane, thereby reducing noise from redundant attributes. This confirms that targeted feature selection not only reduces computational complexity but also strengthens the overall robustness of the classification process, providing a model that can be integrated into clinical workflows as a decision-support tool for flagging high-risk patients, prioritizing further testing, and promoting earlier intervention in thyroid disease management.

Additionally, the study highlights that while RFE contributes most significantly to accuracy improvement, the ADASYN oversampling technique remains essential for enhancing minority-class recognition. By generating synthetic samples in difficult-to-learn regions, ADASYN improves recall, precision, and F1-score, resulting in a model that performs more equitably across both classes and reduces the likelihood of missed diagnoses in imbalanced data settings. The combined effects of feature refinement and class rebalancing produce a harmonized model capable of accurately identifying thyroid disorders even within an imbalanced dataset, offering a practical framework that can be further developed for deployment in hospital information systems or screening programs. Future research should validate this approach on external thyroid datasets from different institutions and populations, systematically compare RFE with other feature-selection techniques (such as LASSO or tree-based embedded methods), and explore alternative SVM kernels as well as comparisons with ensemble and deep learning models to more comprehensively position this framework within the broader landscape of medical diagnostic tools.

REFERENCES

[1] M. S. Salsabilah *et al.*, "HIPOTIROIDSME : Etiologi, Faktor Risiko Dan Tatalaksana Komprehensif," *J. Kesehat. Tambusai*, vol. 5, no. 4, pp. 13211–13218, Dec. 2024.
[2] N. Kadek Mega Suryantini *et al.*, "Gangguan Hormon Tiroid: Hipotiroidisme," *J. Ilmu Kedokt. dan Kesehat.*, vol. 11, no. 6, pp. 1227–1234, Jun. 2024.
[3] R. Wenilia and P. Dokter, "Hypothyroid and Heart Disease," *J. Med. Hutama*, vol. 6, no. 01 Oktober, pp. 4084–4095, Oct. 2024.
[4] S. Andersen *et al.*, "Interpretation of TSH and T4 for diagnosing minor alterations in thyroid function: a comparative analysis of two separate longitudinal cohorts," *Thyroid Res.*, vol. 15, no. 1, pp. 1–8, Dec. 2022.
[5] A. C. Bianco, "Emerging Therapies in Hypothyroidism," *Annu. Rev. Med.*, vol. 75, p. 307, Jan. 2023.
[6] M. D. Ettleson and M. Papaleontiou, "Evaluating health outcomes in the treatment of hypothyroidism," *Front. Endocrinol. (Lausanne).*, vol. 13, p. 1026262, Oct. 2022.
[7] M. Hu *et al.*, "Development and preliminary validation of a machine learning system for thyroid dysfunction diagnosis based on routine laboratory tests," *Commun. Med.*, vol. 2, no. 1, Dec. 2022.
[8] K. V. A. Reddy, S. R. Ambati, Y. S. Rithik Reddy, and A. N. Reddy, "AdaBoost for Parkinson's Disease Detection using Robust Scaler and SFS from Acoustic Features," *Proc. - 1st Int. Conf. Smart Technol. Commun. Robot. STCR 2021*, Oct. 2021.
[9] Alwaliyanto, S. K. Gusti, I. Afrianty, and F. Syafria, "Penerapan Metode ADASYN Dalam Mengatasi Imbalanced Data Untuk Klasifikasi Penyakit Stroke Menggunakan Support Vector Machine," *Bull. Comput. Sci. Res.*, vol. 5, no. 4, pp. 532–541, Jun. 2025.
[10] O. A. M. López, A. M. López, and D. J. Crossa, "Support Vector Machines and Support Vector Regression," *Multivar. Stat. Mach. Learn. Methods Genomic Predict.*, pp. 337–378, Jan. 2022.
[11] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Inf. 2024, Vol. 15, Page 235*, vol. 15, no. 4, p. 235, Apr. 2024.
[12] L. R. Safitri, N. Chamidah, T. Saifudin, M. Firmansyah, and G. T. Alpandi, "Comparison of Logistic

Regression and Support Vector Machine in Predicting Stroke Risk," *Inferensi*, vol. 7, no. 2, pp. 123–128, Jul. 2024.
[13] H. Sundari, M. A. Amrustian, A. Dwi, and P. Wicaksono, "Penerapan Recursive Feature Elimination pada Support Vector Machine untuk Klasifikasi Kanker Payudara," *LEDGER J. Inform. Inf. Technol.*, vol. 3, no. 2, pp. 60–65, Aug. 2024.
[14] M. Yusuf, A. Haq, and S. Rochimah, "Integrating Adaptive Sampling with Ensembles Model for Software Defect Prediction," *Kinet. Game Technol. Inf. Syst. Comput. Network, Comput. Electron. Control*, vol. 10, no. 2, pp. 229–240, May 2025.
[15] A. Fitri, I. Afrianty, E. Budianita, and S. K. Gusti, "Implementation of Feature Selection Information Gain in Support Vector Machine Method for Stroke Disease Classification," *Bull. Informatics Data Sci.*, vol. 4, no. 1, pp. 22–33, May 2025.
[16] A. Raza, F. Eid, E. C. Montero, I. D. Noya, and I. Ashraf, "Enhanced interpretable thyroid disease diagnosis by leveraging synthetic oversampling and machine learning models," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, pp. 1–18, Dec. 2024.
[17] J. H. Chen, Y. Q. Zhang, T. T. Zhu, Q. Zhang, A. X. Zhao, and Y. Huang, "Applying machine-learning models to differentiate benign and malignant thyroid nodules classified as C-TIRADS 4 based on 2D-ultrasound combined with five contrast-enhanced ultrasound key frames," *Front. Endocrinol. (Lausanne).*, vol. 15, p. 1299686, Apr. 2024.
[18] L. B. V De Amorim, G. D. C. Cavalcanti, and R. M. O. Cruz, "The choice of scaling technique matters for classification performance."
[19] M. Koch, J. Xia, and C. A. Ramezan, "Transferability of Recursive Feature Elimination (RFE)-Derived Feature Sets for Support Vector Machine Land Cover Classification," *Remote Sens. 2022, Vol. 14, Page 6218*, vol. 14, no. 24, p. 6218, Dec. 2022.
[20] S. K. Lin, H. Hsiu, H. S. Chen, and C. J. Yang, "Classification of patients with Alzheimer's disease using the arterial pulse spectrum and a multilayer-perceptron analysis," *Sci. Rep.*, vol. 11, no. 1, p. 8882, Apr. 2021.
[21] Y. S. Kim, M. K. Kim, N. Fu, J. Liu, J. Wang, and J. Srebric, "Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models," *Sustain. Cities Soc.*, vol. 118, p. 105570, Jan. 2025.
[22] T. Taryadi, E. Yunianto, and K. Kasmari, "Diagnostik Penyakit Ginjal Kronis Menggunakan Model Klasifikasi Support Vector Machine," *IC Tech Maj. Ilm.*, vol. 19, no. 1, pp. 39–44, Apr. 2024.
[23] W. Li, S. Zhu, Z. Li, and H. Wang, "Kernel-Based Enhanced Oversampling Method for Imbalanced Classification," Apr. 2025.
[24] A. Demircioğlu, "The effect of feature normalization methods in radiomics," *Insights Imaging*, vol. 15, no. 1, p. 2, Dec. 2024.
[25] M. Mujahid *et al.*, "Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering," *J. Big Data 2024 111*, vol. 11, no. 1, pp. 87-, Jun. 2024.
[26] Y. Zhu, K. Liu, M. Gu, K. Zhang, and G. Hu, "Image recognition method of cashmere and wool based on SVM-RFE selection with three types of features," *Autex Res. J.*, vol. 25, no. 1, Jan. 2025.
[27] L. van Bemmel, W. Harmsen, C. Cucchiarini, and H. Strik, "Automatic Selection of the Most Characterizing Features for Detecting COPD in Speech," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12997 LNAI, pp. 737–748, Sep. 2021.
[28] J. Kalpathy-Cramer, J. B. Patel, C. Bridge, and K. Chang, "Basic Artificial Intelligence Techniques: Evaluation of Artificial Intelligence Performance," *Radiol. Clin. North Am.*, vol. 59, no. 6, pp. 941–954, Nov. 2021.
[29] S. Sathyanarayanan and B. R. Tantri, "Confusion Matrix-Based Performance Evaluation Metrics," *African J. Biomed. Res.*, vol. 27, no. 4S, pp. 4023–4031, Nov. 2024.
[30] Y. Yang, H. A. Khorshidi, and U. Aickelin, "A review on over-sampling techniques in classification of multi-class imbalanced datasets: insights for medical problems," *Front. Digit. Heal.*, vol. 6, p. 1430245, 2024.

KOMUNITAS
DOSEN INDONESIA

Editor Decision
KDI/992/LoA-Invoice/2025
10/12/2025

Dears Mr/Mrs/Ms. **Citra Wulandari, Iis Afrianty , Elvia Budianita, Siska Kurnia Gusti**
Thank you for your trust in our services.

With this confirmation, we have decided on your submission to **bit-Tech ISSN. 2622-271X (Print) & 2622-2728 (Online)**, **Thyroid Disease Classification Using Support Vector Machine and Recursive Feature Elimination Method** has been **Accepted**, for Vol 8 No 2 according to the results of the reviewer's reading notes.

The article will be uploaded and published online with an estimated time is **10/12/2025** softcopy (e-version).

Thank you for your attention and cooperation.

Sincerely,

Komunitas Dosen Indonesia

payment at least 3 days after the article is received: 12/12/2025
Please confirm payment via wa: 081807834703
