APPLICATION OF BACKPROPAGATION NEURAL NETWORK USING RANDOM OVERSAMPLING AND ROBUST SCALER FOR CLASSIFICATION THYROID

TUGAS AKHIR

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

Oleh
UMMY AGUSTINA PUTRI
NIM. 12150124340

UIN SUSKA RIAU

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

LEMBAR PERSETUJUAN

APPLICATION OF BACKPROPAGATION NEURAL NETWORK USING RANDOM OVERSAMPLING AND ROBUST SCALER FOR CLASSIFICATION THYROID

LAPORAN TUGAS AKHIR MAHASISWA
JURUSAN TEKNIK INFORMATIKA
UIN SUSKA RIAU

TUGAS AKHIR

Oleh

UMMY AGUSTINA PUTRI
NIM. 12150124340

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 8 Januari 2026

Pembimbing I

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

LEMBAR PENGESAHAN

APPLICATION OF BACKPROPAGATION NEURAL NETWORK USING RANDOM OVERSAMPLING AND ROBUST SCALER FOR CLASSIFICATION THYROID

LAPORAN TUGAS AKHIR MAHASISWA
JURUSAN TEKNIK INFORMATIKA
UIN SUSKA RIAU

Oleh

UMMY AGUSTINA PUTRI
NIM. 12150124340

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, 8 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	: Dr. Alwis Nazir, M. Kom
Pembimbing I	: Iis Afrianty, S.T., M.Sc.
Penguji I	: Elvia Budianita, S.T., M.Cs.
Penguji II	: Fadhilah Syafira,S.T.,M.Kom

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama	: Ummy Agustina Putri
NIM	: 12150124340
Tempat/Tgl Lahir	: Tanah Merah, 14 Agustus 2003
Fakultas	: Sains dan Teknologi
Prodi	: Teknik Informatika
Judul Skripsi	: Application of Backpropagation Neural Network Using Random Oversampling and Robust Scaler for Classification Thyroid

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

UMMY AGUSTINA PUTRI

NIM. 12150124340

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Tugas Akhir ini dengan judul "Application Of Backpropagation Neural Network Using Random Oversampling And Robust Scaler For Classification Thyroid" 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 2026
Yang membuat pernyataan,

UMMYAGUSTINA PUTRI
NIM.12150124340

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 PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

Alhamdulillahi robbil'alamin..

Segala puji bagi Allah SWT. Dalam setiap langkah yang dilalui, penulis menyadari bahwa tidak ada kekuatan selain pertolongan-Nya. Atas izin, kasih sayang, dan kehendak-Nya, penulis diberi kemampuan untuk bertahan, belajar, dan terus melangkah hingga karya tulis ini dapat diselesaikan. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad SAW, yang menjadi cahaya penuntun dalam menghadapi perjalanan hidup.

Dengan segala cinta dan ketulusan, karya tulis ini penulis persembahkan kepada:

Orang tua tercinta, *Bapak Abd. Rasyid dan Mamak Wanti,* adalah rumah pertama dan tempat pulang terbaik bagi penulis. Dengan cinta yang tak pernah habis, kesabaran yang tak pernah lelah, serta pengorbanan yang sering kali tidak terlihat, Bapak dan Mamak telah berhasil mendidik, membimbing, dan menguatkan penulis hingga Ummy tumbuh dan menjadi pribadi seperti saat ini. Dari Bapak dan Mamak, penulis belajar tentang arti tanggung jawab, keikhlasan, dan keteguhan dalam menjalani kehidupan. Setiap langkah yang penulis tempuh selalu diiringi doa yang tak pernah putus, pengorbanan yang tak terhitung, serta kasih sayang yang tak pernah berkurang. Terima kasih atas cinta yang tulus, dukungan yang tak ternilai, dan kepercayaan yang senantiasa diberikan hingga penulis mampu menyelesaikan perjalanan panjang dalam menempuh pendidikan ini. Tanpa doa, restu, dan pengorbanan Bapak dan Mamak, penulis tidak akan pernah berada di titik ini.

Saudara kandung penulis, *Abang Edward Dede Ashari,* senantiasa hadir dengan kehangatan, perhatian, dan dukungan melalui canda sederhana yang sarat makna. Di balik keusilannya, tersimpan kepedulian serta nasihat yang menguatkan penulis untuk tetap teguh, sabar, dan ikhlas menjalani setiap proses. Penulis juga menyampaikan terima kasih yang setulusnya kepada seluruh keluarga, khususnya *Bunda Sutra Dewi*, serta keluarga besar yang tidak dapat disebutkan satu per satu, atas doa, kasih sayang, dan perhatian yang senantiasa menjadi sumber kekuatan dan semangat.

Sahabat terbaiku saudari tidak sedarah, *Ade Irna Titania, Nazhifah Al-Adawiyah Tenrisau, Nur Aziza Putri, Ranti Aurelia Putri, dan Yohana Putri,* sahabat satu perjuangan menuju gelar sarjana—mulai dari saling menyemangati, mengeluh bersama, sampai tertawa di tengah lelah. Terima kasih atas kebersamaan, dukungan, dan semangat yang selalu hadir. Semoga Allah SWT memberkahi setiap langkah kita dan menjaga persahabatan ini hingga kapan pun.

Seluruh dosen Prodi Teknik Informatika, Khususnya *Ibu Pembimbing Iis Afrianty, S.T., M.Sc.,* serta *Ibu Elvia Budianita, S.T., M.Sc*. dan *Ibu Fadhilah Syafria, S.T., M.Kom.* selaku dosen penguji, penulis menyampaikan rasa terima kasih yang sebesar-besarnya atas bimbingan, arahan, kritik, dan masukan yang sangat berarti selama proses penyusunan karya ini. Waktu, tenaga, dan perhatian yang diberikan menjadi bagian penting hingga karya pertama penulis dapat diselesaikan dengan baik dan tepat waktu. Semoga segala kebaikan tersebut bernilai ibadah dan menjadi amal jariyah yang pahalanya terus mengalir.

Saudari seperjuangan selama masa perkuliahan, *Citra Wulandari* dan *Shella Sabila*, meski tidak terikat oleh hubungan darah, kebersamaan yang terjalin terasa layaknya keluarga. Terima kasih atas kebersamaan yang tak ternilai, tawa yang menguatkan di tengah lelah, serta kehadiran yang setia menemani setiap proses dan perjuangan. Dukungan, doa, dan semangat yang kalian berikan menjadi kenangan berharga dalam perjalanan penulis. Semoga kelak Allah SWT mempertemukan kita kembali dalam keadaan terbaik, di waktu dan tempat yang penuh keberkahan. **Seluruh teman-teman Jurusan Teknik Informatika Angkatan 2021** terima kasih atas kebersamaan, perhatian, dan dukungan yang tulus selama menjalani masa perkuliahan. Setiap bantuan, tawa, dan kebersamaan yang terjalin menjadi kenangan yang akan selalu dikenang dalam perjalanan penulis.

Seseorang yang istimewa, tersayang, dan sangat berarti, *Shafiq Zahidan, S.T.*, hadir bukan sekedar sebagai bagian dari perjalanan ini, melainkan hadir dengan cinta yang penuh ketulusan dan bermakna. Pertemuan yang tak terduga berkembang menjadi proses saling memahami dan menguatkan. Kehadiranmu dalam penyusunan karya ini menjadi bagian penting yang mengiringi hingga karya ini terselesaikan. Hadirmu menjadi rumah bagi penulis; tempat yang membahagiakan setelah lelah, tempat bercerita saat ragu, dan tempat penulis merasa di mengerti dengan segala bentuk dukungan, pengertian, canda tawa, tenaga, waktu, dan cinta. serta penguat sekaligus penenang di setiap langkah perjalanan penulis. Terima kasih atas perhatian, kepedulian, dan kasih sayang yang diberikan dengan tulus, kasih sayang yang sering kali tanpa diminta dan tanpa disadari. serta motivasi untuk terus belajar dan melangkah bersama. Semoga Allah SWT senantiasa melimpahkan keberkahan, kebahagiaan, dan kebaikan dalam hidupmu, serta menjaga ikatan dan kasih sayang yang telah terjalin di antara kita.

Diri saya sendiri, *Ummy Agustina Putri,* anak perempuan bungsu yang setiap harinya belajar menguatkan diri untuk tetap yakin, bersemangat, dan terus melangkah menuju kebahagiaan. Terima kasih telah bertahan, berjuang, berani, dan tidak menyerah di tengah berbagai rintangan yang dihadapi selama proses karya ini. Setiap lelah, air mata, dan keraguan telah menjadi bagian dari perjalanan yang membentuk diri ini menjadi versi yang lebih kuat dan lebih baik. Karya ini merupakan bukti bahwa dengan kesabaran, ketekunan, dan kepercayaan pada diri sendiri, segala tantangan dapat dilalui dengan baik. Semoga langkah ke depan selalu dipenuhi kemudahan, keberanian , dan kebahagiaan, serta mampu menjalani kehidupan dengan penuh rasa syukur.

Karya ini dipersembahkan sebagai pengingat bahwa setiap perjuangan memiliki makna. Bahwa jatuh, lelah, dan ragu bukanlah tanda kegagalan, melainkan bagian dari perjalanan menuju versi diri yang lebih kuat, dan setiap langkah akan dipermudah ketika disertai doa serta usaha yang sungguh-sungguh.

QS. Al-Insyirah (94): 5–6

"Karena sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan."

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

Application of Backpropagation Neural Network Using Random Oversampling and Robust Scaler for Classification Thyroid

Emmy Agustina Putri[1)], Iis Afrianty[2)*], Elvia Budianita[3)], Fadhilah Syafria[4)]
[1)2)3)] *Informatics Engineering, Science and Technology, Universitas Islam Negeri Sultan Syarif Karim Riau, Pekanbaru, Indonesia*
[1)]12150124340@students.uin-suska.ac.id
[2)]iis.afrianty@uin-suska.ac.id
[3)]elvia.budianita@uin-suska.ac.id
[4)]fadhilah.syafria@uin-suska.ac.id

Article history:

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

Keywords:

Artificial Neural Network;
Backpropagation;
Thyroid Classification;
Random Over Sampling;
RobustScaler;

Abstract

Thyroid disease is a fairly common endocrine disorder that requires rapid and accurate diagnosis so that patients can receive appropriate treatment. This study was conducted to improve the system's ability to classify thyroid disease by utilizing data preprocessing techniques with RobustScaler and Random Over Sampling (ROS), as well as the Backpropagation Neural Network (BPNN) algorithm. The research dataset consisted of 3,771 patient data with 25 clinical attributes describing the condition and function of the thyroid. The data preprocessing process involved data selection, data cleaning, and data transformation using RobustScaler so that each feature had a more stable scale and was not affected by extreme values. The class imbalance problem was overcome using ROS so that the amount of data increased to 6,834 samples and the class distribution became more balanced. The Backpropagation Neural Network algorithm was applied in model training by testing various variations in the number of neurons in the hidden layer (38 and 49) and learning rate (0.01 and 0.001). Training was conducted for 5,000 and 10,000 epochs. Evaluation was performed using the 10-Fold Cross Validation method to obtain more consistent results. The results of the study show that the model is capable of achieving very high accuracy, up to 99.85%, on several parameters. The results show that proper data processing and appropriate parameter selection greatly affect model performance. Overall, the use of RobustScaler and ROS has been proven to significantly improve the accuracy of thyroid disease classification.

I. INTRODUCTION

The thyroid is a key endocrine organ that regulates metabolism through the hormones thyroxine (T4) and triiodothyronine (T3), which are stimulated by TSH from the pituitary gland [1] [2] [3]. Barious diseases that can occur due to thyroid gland dysfunction include: Hypothyroidism, Hyperthyroidism, Goiter, Thyroiditis, Graves' Disease, Thyroid Nodules, and Thyroid Cancer [4] [5]. In recent years, the incidence of thyroid disorders has risen, particularly hypothyroidism, hyperthyroidism, and thyroid nodules. In Australia, thyroid tests covered by Medicare increased from 7.37 million in 2014 to 9.89 million in 2023 [6]. Similarly, in Indonesia, around 12.4 million people have hypothyroidism, with only 1.9% receiving treatment, and 13.2 million have hyperthyroidism, with just 6.2% treated [7] indicating that a large proportion of patients remain undiagnosed or untreated despite the substantial disease burden. This condition shows that, even though the number of cases is high, the treatment rate is still low, so the epidemiological impact of thyroid disease has not been optimally addressed and remains an important public health challenge [8]. Symptoms such as fatigue, cold intolerance, and weight gain are common but often overlooked, making early detection challenging [9][10].

ISSN 2622-2728 (online) 2622-271X (print)
doi: 10.32877/bt.v8i2.3455

The main challenge in diagnosing thyroid disease is that symptoms are often mild and nonspecific [11]. They are often misinterpreted as common complaints, such as fatigue, weight gain, or intolerance to cold temperatures [12]. Consequently, many cases are not detected early, and clinical decisions may rely heavily on subjective judgement rather than systematic risk assessment. Laboratory tests using TSH, T3, TT4, and T4U parameters are important indicators in determining thyroid function, but interpreting test results often requires complex analysis, especially when the data contains extreme or unbalanced values [13]. Under these conditions, traditional manual assessment becomes prone to misclassification and bias, so there is a strong need for data-driven decision-support models that can assist clinicians in interpreting heterogeneous laboratory and clinical data more objectively.

Previous studies have explored thyroid disease classification using various datasets and methods. One study [14] used a Kaggle thyroid dataset with 3,772 samples and 30 features, including patient characteristics, medical conditions, and hormone levels (TSH, T3, TT4, T4U, FTI), for classification. Due to class imbalance and mixed data types, encoding and SMOTE resampling were applied to balance the dataset. The Light Gradient Boosting Machine (LGBM) achieved the highest accuracy of 96%, demonstrating the effectiveness of preprocessing and model selection. However, this performance also indicates that there is still room for improvement, particularly in handling imbalance and extracting nonlinear relationships between laboratory parameters and disease status. Furthermore, research by [15] compared various machine learning algorithms using the *sick-euthyroid* dataset from UCI, which contains 3,162 thyroid disease data, with a highly imbalanced class distribution, namely 2,870 negative data and 292 positive data. The study did not use oversampling or other balancing techniques, relying solely on metrics such as F1-score to assess the model's ability to recognize minority classes. Testing of several machine learning algorithms showed significant variations in performance as follows: Random Forest (RF) achieved the highest accuracy of 97.67%, followed by Support Vector Machine (SVM) with an accuracy of 96.05%, and Decision Tree (DT) with 95.35%. Meanwhile, Naïve Bayes (NB) recorded an accuracy of 92.44%, while Logistic Regression (LR) showed the lowest performance with an accuracy of 90.70%. These results confirm that advanced models such as RF and SVM can achieve high accuracy, but they also highlight that ignoring class imbalance may mask poor detection of rare yet clinically important positive cases.

Based on previous research, thyroid disease datasets are often imbalanced, where the number of negative classes is much greater than positive classes, causing models to be biased towards the majority class[16]. To overcome this, Random Over Sampling (ROS) is used, which is a technique that randomly adds samples from the minority class to make the proportions more balanced [17]. A study by **c**ompared Random Over Sampling (ROS) and Random Under Sampling (RUS) using a stroke dataset with 3,889 non-stroke patients (majority class) and only 198 stroke patients (minority class) [18]. The results showed that when the data was balanced using (ROS), the model achieved 95% accuracy. Conversely, when using RUS, the accuracy dropped to 76%. This proves that ROS is more effective than RUS in handling data imbalance. Another study, processed 9,172 unbalanced thyroid data and tested several resampling methods [19]. Random Oversampling produced the highest accuracy, namely 98.5%, thus proving to be effective in balancing data by randomly adding minority class samples through (ROS). The model can improve its sensitivity to that class without eliminating the majority class, so that the classification performance becomes more balanced [20].

In addition to addressing class imbalance in the thyroid disease dataset, another important step in the preprocessing process is to make the feature scale consistent across each feature through normalization techniques [19]. Normalization is necessary because medical variables such as TSH, T3, TT4, and T4U levels have different value ranges and may contain extreme values (outliers) [20]. Outliers can interfere with statistical calculations such as mean and standard deviation, thereby significantly affecting data analysis results [21]. To overcome this limitation, this study uses the RobustScaler method, which utilizes the median and interquartile range (IQR) so that the normalization results are more stable against extreme values compared to other methods such as StandardScaler or MinMaxScaler, and better reflect the real distribution of clinical laboratory data.

Furthermore, compares the performance of the ID3 and Artificial Neural Network (ANN) algorithms in classifying thyroid disease [22]. Testing was conducted using a data division scheme based on percentage and the 5-fold cross validation method. The results show that ANN provides better performance, with an accuracy of 90% in data division by percentage and 95% in 5-fold cross validation. Meanwhile, ID3 only achieved an accuracy of 88% and 93% in the same scheme. Based on these results, it can be concluded that ANN is superior in terms of accuracy with 5-fold cross validation, and its performance has the potential to improve further if 10-fold cross validation is applied.

In this study, an Artificial Neural Network (ANN) is implemented using the Backpropagation Neural Network (BPNN) method. BPNN uses supervised learning to update weights based on the difference between predicted and target outputs through backpropagation [23]. Each output neuron maps data to a specific class, allowing the model to learn complex nonlinear patterns. According to [24] BPNN can capture more complex patterns than classical algorithms, making it effective for classifying thyroid diseases. The study [25] compared Decision Tree C4.5 and BPNN in classifying smartphone prices with 3,000 data points and 21 attributes. C4.5 achieved 83.75% accuracy, while BPNN with a 12-15-8-1 architecture, 500 epochs, a learning rate of 0.01, and momentum of 0.1 achieved 96.65% accuracy, showing BPNN's superiority. In this research, clinical symptoms and medical parameters are used as inputs, with positive and negative diagnosis labels as outputs. Random Over Sampling (ROS) is applied to handle class imbalance, and Robust Scaler normalizes features to reduce outlier effects. Model performance is evaluated using 10-fold cross-validation to ensure reliable results. This approach aims to improve accuracy, support early detection, and enhance thyroid disease prevention and control.

II. Related Works/Literature Review

Thyroid disease is a disorder of the thyroid gland, a butterfly-shaped endocrine gland located at the front of the neck [26]. This gland produces the hormones T4 and T3, which play a role in regulating metabolism, growth, and body systems [12]. This study [27] Pegular testing of TSH and FT4 levels is highly recommended to facilitate the diagnosis and treatment of thyroid disease. Several factors such as genetics, gender, age, exposure to chemicals, stress, and iodine intake also influence the risk of thyroid disorders [10]. In medical data processing, the pre-processing stage plays a crucial role in enhancing model accuracy. Previous research [28] shows that the application of Robust Scaler can significantly improve model performance, with the F1-Score increasing from 0.68 (without scaling) to 0.83 (with *Robust Scaler)*, and accuracy increasing from 0.61 to 0.82. Research [19] shows that the choice of normalization technique has a major impact on classification performance. Five methods were tested: StandardScaler, MinMaxScaler, MaxAbsScaler, RobustScaler, and QuantileTransformer. The results show that RobustScaler is the most superior. This method uses the median and interquartile range *(IQR)* so that it is resistant to *outliers* and maintains the stability of the data distribution.Class imbalance is also a common problem in disease diagnosis data. The importance of handling class imbalance (imbalanced data) in disease diagnosis, Research conducted by [29] They processed 9,172 imbalanced thyroid records and evaluated several resampling methods. Random oversampling produced the highest accuracy, at 98.5%, proving to be the most effective in balancing data and improving classification results; this evidence reinforces the importance of explicit imbalance handling in medical classification problems.

Additionally, several studies indicate that Backpropagation Neural Network (BPNN) is a method capable of producing high accuracy due to its ability to deeply learn data patterns [30] 's research tested the Backpropagation Neural Network method using 500 patient data with learning rates varying from 0.0001, 0.001, 0.01, and 0.1, as well as 9–17 hidden neurons. In the 90:10 scheme, the highest accuracy reached an average of 98.42%, with some configurations even reaching 100%, particularly with learning rates of 0.001–0.1 and 17 neurons. This study [31] Research Classification of stroke using BPNN showed an accuracy of 96.14%, with medical records data covering age, hypertension, and blood sugar levels. The BPNN model proved effective in distinguishing between stroke and non-stroke patients [32], demonstrating its suitability as a decision-support tool in critical diagnostic contexts. Another line of work on Parkinson's disease classification using Backpropagation Neural Network (BPNN) was able to achieve an accuracy of up to 99.6%. Research in this area generally uses datasets of 700–1,300 samples, such as the UCI or Kaggle datasets, which contain approximately 756 voice recordings and clinical features of Parkinson's patients. Research [33] uses BPNN with 9 neurons in the hidden layer, 50 iterations, and a learning rate of 0.3. This model achieved an accuracy of 99.49%, with a True Positive Rate (TPR) of 99.5% and a False Positive Rate (FPR) of only 0.08%, indicating that BPNN can achieve very high sensitivity while keeping false alarms at a minimal level. The study [34] shows that the Backpropagation Neural Network (BPNN) method is superior to Case-Based Reasoning (CBR). From 120 data tested using 3-fold, 4-fold, and 5-fold cross validation, BPNN obtained an average accuracy of 95.07%, while CBR only reached 94.70%. This difference shows that BPNN is more capable of learning complex patterns and producing more stable predictions.

III. Methods

This study follows a systematic and interconnected workflow in the application of BPNN with Random Oversampling and Robust Scaler for thyroid disease classification. All stages, from data processing to model evaluation, are shown in Figure 1.

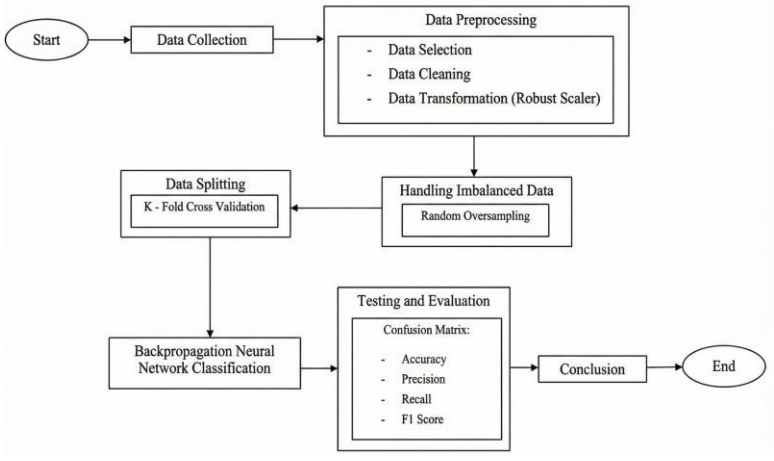

Fig. 1 Research Stages

A. Data Collection

The data used in this study is secondary data sourced from the *Kaggle* platform, via the link https://www.kaggle.com/datasets/yasserhessein/thyroid-disease-data-set. The *Thyroid Disease* dataset consists of 3,771 data points with 25 attributes. The attributes used in this study can be seen in Table 1.

TABLE 1
DATASET ATTRIBUTES

Attribute	Description	Attribute	Description
Age	Patient age	Tumor	Having a tumor
Sex	Patient's gender	Hypopituitary	Pituitary gland disorder
On_thyroxine	Currently taking thyroxine hormone medication	Psych	Psychological disorders
Query_on_thyroxine	History of thyroxine medication use	TSH_measured	TSH measurement status
On_antithyroid_medication	Currently taking antithyroid medication	TSH	TSH level value
Sick	Currently ill	T3_measured	T3 measurement status
Pregnant	Currently pregnant	T3	T3 level value
Thyroid surgery	Has had thyroid surgery	TT4_measured	TT4 measurement status
I131 treatment	Have undergone radioactive iodine therapy	TT4	TT4 level value
Query_hypothyroid	Suspected hypothyroidism	T4U_measured	T4U measurement status
Query_hyperthyroid	Suspected hyperthyroidism	T4U	T4U level value
Lithium	Taking lithium	FTI	Free Thyroxine Index Value
Goiter	Goiter or thyroid enlargement		

The original dataset used in this study consists of two classes, namely class 0 for negative patients (without thyroid disease) and class 1 for positive patients (with thyroid disease), as shown in Table 2.

TABLE 2
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
...	...	...	...	...	...	...	...	...
3770	72	1	...	0.7	82	0.94	87	0
3771	64	0	...	1	99	1.07	92	0

B. Data Preprocessing

1) Data Selection

This stage involves selecting the data needed for classification. There are 291 people with thyroid disease and 3,480 people without thyroid disease. All data and attributes are used for the next modeling stage.

2) Data Cleaning

At this stage, *data cleaning* was performed to ensure data quality before modeling. The steps taken included handling *missing values*, removing 63 duplicate data, and checking outliers on continuous features such as *Age, TSH, TT4, T4U,* and *FTI*. Descriptive statistical analysis was used to see the data distribution and identify possible extreme values that could affect the model results. Outlier detection in the dataset is shown in Table 3.

TABLE 3
DETECTION OF OUTLIERS IN THE DATASET

Statistics	Age	TSH	TT4	T4U	FTI
count	3708	3708	3708	3708	3708
mean	51.746962	5.088809	108.325450	0.994954	110.481499
std	19.007393	23.490882	34.788608	0.186734	31.618426
min	1.000000	0.005000	2.000000	0.25	2.000000
25	36.000000	0.580000	88.75	0.89	94.000000
50	54.0000	1.550000	105.00000	0.99	110,000,000
75	67,000,000	3,500,000	123,000,000	1,070,000	122,000,000
max	94,000,000	530,000,000	430,000,000	2,320,000	395,000

Table 3 shows that the TSH, TT4, and FTI features have maximum values far above the Q3 value, so they potentially contain outliers. Conversely, Age and T4U have more stable distributions without extreme deviations. This identification is important to prevent bias due to incomplete or inconsistent data, so that data quality improves and prediction models can work more accurately.

3) Data Transformation

At this stage, data transformation is performed to stabilize the scale of features with large ranges and containing outliers. Based on Table 2, several features have scale differences and extreme values that can reduce model performance. To overcome this, RobustScaler is used, which is a normalization method that reduces the median value and then divides it by the Interquartile Range (IQR) so that the influence of outliers can be minimized. The normalization process follows the robustscaler formula using equations 1 and 2:

$$IQR(X) = Q_{\{3\}} - Q_{\{1\}} \quad (1)$$

The Interquartile Range (IQR) is the difference between the third quartile (Q3) and the first quartile (Q1), representing the middle 50% of the data and reducing the effect of outliers.

$$X' = \frac{(X - Median(X))}{IQR(X)} \quad (2)$$

X′ is the transformed value, and X is the original data point. The median is used as a central reference because it is resistant to outliers, while the Interquartile Range $IQR(X)$ acts as the scaling factor to produce a more uniform distribution.

C. Data Imbalance Handling

At this stage, class balancing was performed because the number of samples in each class was not proportional, namely class 0 had 3,417 data while class 1 only had 291 data. This imbalance had the potential to make the model biased towards the majority class. Therefore, the Random Oversampling technique was used, which involved adding random samples to the minority class until the number was equal to that of the majority class. The process of balancing data between the majority and minority classes in this study is visualized through the Random Oversampling (ROS) flow, as shown in Figure 2.

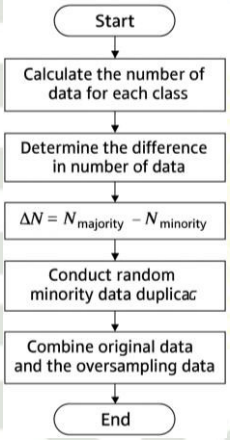

Fig.1Random Oversampling (ROS) process flow

Calculating the Amount of Data in Each Class : The oversampling process begins to address class imbalance in the training data. The first step is to enter the amount of training data with the amount of data from each class. Calculate which class has more data and which has less.

Determining the Difference in Data Amount : Calculate the data difference using the formula (3):

$$\Delta N = N_{mayoritas} - N_{minoritas} (3)$$

ΔN represents the number of new data points that need to be added as a result of oversampling. $N_{mayoritas}$ refers to the amount of data in the class with the largest number of samples, while $N_{minoritas}$ indicates the amount of data in the class with the fewest samples.

Perform Random Minority Data Duplication :The iteration is performed ΔN times. For each iteration i, select one data point from the minority class at random (*random sampling with replacement*). Duplicate that data point. Finally, add it to the minority dataset.Iteration formula (4):

$$Xminoritas_{baru(i)} = Random(X_{minoritas}) \quad (4)$$

$Xminoritas_{baru(i)}$refers to the new data produced during iteration iii through random replication. The function $Random(X_{minoritas})$ generates these values by selecting one sample from the minority set with equal probability. Meanwhile, $X_{minoritas:}$ represents the full collection of original data belonging to the minority class from which the selections are made.

Explanation:

Combine Original Data and Oversampling Results: The new dataset is formed as follows formula (5):

$$D_{baru} = D_{mayoritas} \cup D_{minoritas_{asli}} \cup D_{minoritas_{baru}} (5)$$

D_{baru} is the final dataset after applying ROS. $D_{mayoritas}$ represents the original majority class data, while $D_{minoritas_{asli}}$ refers to the minority class data before oversampling. $D_{minoritas_{baru}}$ denotes the minority data generated through random oversampling.

D. Data Division

K-Fold Cross Validation is a model evaluation technique that divides data into several multiples to obtain more objective and stable results[35] . In *K-Fold Cross Validation*, the dataset is divided into 'k' multiples of relatively equal size. Each multiple takes turns acting as test data, while the other multiples are used as training data. This process occurs for 'k' iterations so that each part of the data has the opportunity to become test data[36] In research by[35] , it is explained that *K-Fold Cross Validation* has the advantage of providing more stable and accurate model evaluation results through the average of k iterations so that the results are more reliable because the model is tested with various data combinations, and the risk *of overfitting* can be minimized. In this study, the *k-fold cross-validation* technique was used with a data division of 10 k.

E. Backpropagation Neural Network

The Backpropagation Neural Network (BPNN) is a learning algorithm commonly used in multi-layer artificial neural networks (multi-layer perceptrons)[37]. Learning occurs by updating network weights based on the error between the actual output and the target [38]. BPNN has three main layers: input, hidden, and output. The hidden layer is crucial because it captures non-linear representations of the data, improving classification accuracy [39] . Common activation functions include sigmoid, tansig (tanh sigmoid), and ReLU. During the forward pass, input data is processed through the hidden neurons using weights and activation functions to produce outputs[40] In the backward pass, errors are calculated and propagated back to adjust the weights, which are then updated to reduce the error. This process repeats until the desired accuracy is reached or the iteration limit is met. According to[41] BPNN training generally includes initialization, feedforward, backpropagation, and weight updates. The detailed stages are as follows:

1) Parameter Initialization In the initial stage, weights and biases are randomly initialized with small values. In addition, the learning rate, maximum number of epochs, and target error are set as training references.

Phase I: Feedforward Propagation

2) Each neuron in the input layer receives input datax_i $(x_1, x_2, \ldots\ldots x_n)$. This input value is then forwarded to the hidden layer through the weights that connect it.
3) Each neuron in the hidden layer is marked as a $z_j(j = 1,2, \ldots\ldots, p)$, calculating the linear combination of all incoming inputs, plus the bias, using the formula in Equation 6.

$$z_{in_j} = v_{0j} + \sum_{i=1}^{n} x_i v_{ij} \quad (6)$$

x_i represents the input value to neuron i. v_{ij} is the weight connecting input i to hidden neuron j, and $v0j$ denotes the bias of neuron j.

Next, the values ofz_{in_j} are passed into an activity function, such as sigmoid. Equation 7.

$$z_j = \frac{1}{1 + e^{-z_{in_j}}} \quad (7)$$

4) Each neuron in the output layery_k $(k = 1,2, \ldots, m)$ receives signals from the hidden layer, which are then summed after being multiplied by their respective weights. Equation 8.

$$z_j = \frac{1}{1 + e^{-z_{in_j}}} \quad (7)$$

wjk is the weight connecting hidden neuron$j$ to output neuron k, and w_{0k} represents the bias of the k-th output neuron.

The value ofy_{ink} is then passed through the activation function. Equation 9.

$$y_k = \frac{1}{1 + e^{-y_{ink}}} \quad (9)$$

Phase II: Backpropagation

5) Each *output* unit, marked asy_k $(k = 1,2, \ldots, m)$, where the network outputy_k will be compared with the corresponding target pattern, so that the error can be calculated as the basis for evaluating network performance. Equation 10.

$$\delta_k = (t_k - y_k)\, y_k\, (1 - y_k) \quad (10)$$

The next step is to update the values through the calculation of the weightsWjk and the bias, Equation 11 and Equation 12.

$$\Delta w_{jk} = \alpha \cdot \delta_k \cdot z_j \quad (11)$$
$$\Delta w_{0k} = \alpha \cdot \delta_k \quad (12)$$

α is the learning rate, and z_j is the output of hidden neuron j.

6) Each neuron in the hidden layer, denoted by$(j = 1,2,3, zj, \ldots, p)$, , calculates the delta input by summing all the signals it receives. Equation 13.

$$\delta_{inj} = \sum_{k=1}^{m} \delta_k w_{ij} \quad (13)$$

Then, the delta value for the hidden neuron $-j$ is Equation 14.

$$\delta_j = \delta_{inj} \cdot z_j(1 - z_j) \quad (14)$$

After that, the weight and bias changes for the hidden layer are calculated as Equation 15 and Equation 16.

$$\Delta v_{ij} = \alpha \cdot \delta_j \cdot x_i \quad (15)$$

$$\Delta v_{0j} = \alpha \cdot \delta_j \quad (16)$$

Phase III: Weight Modification

7) Each *hidden* unit $zj(j = 1,2,3,\ldots,p)$ will also undergo updates to its bias and weights with (i =1,2,3,…,n). Equation 17 and Equation 18.
Hidden layer

$$v_{ij}(baru) = v_{ij}\,(lama) + \Delta v_{ij} \quad (17)$$

$$v_{0j}(baru) = v_{ij}\,(lama) + \Delta v_{ij} \quad (18)$$

Each neuron in the output layeryk $(k = 1,2,3,\ldots,m)$ will be updated, both in terms of the bias value yk and the weights connected to it$(j = 1,2,3,\ldots,p)$. Equation 19 and Equation 20.
Output layer

$$w_{jk}(baru) = w_{jk}\,(lama) + \Delta w_{jk} \quad (19)$$

$$w_{0k}(baru) = w_{0k}\,(lama) + \Delta w_{0k} \quad (20)$$

8) After one training cycle (feedforward → backpropagation → weight modification), calculate the total error. If the error is already smaller than the target error or the maximum number of epochs has been reached, then training is stopped. If not, the process continues by returning to Step 2.

F. Testing and Evaluation

Testing is a crucial stage in research to ensure that the designed system functions correctly both technically and structurally. In this study, the implementation focused on applying a Backpropagation Neural Network combined with Random Oversampling and Robust Scaler to balance the dataset and enhance classification performance. The implementation was carried out using Python on Google Colab, allowing the model to be trained and tested effectively and efficiently, resulting in more accurate thyroid disease classification. The evaluation stage was conducted to measure the model's performance. One commonly used method is the Confusion Matrix, which compares the model's predictions with the actual labels to clearly show classification accuracy. Figure 3 presents the Confusion Matrix used to evaluate the model's classification performance.

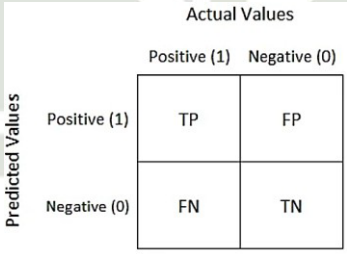

Figure2Confusion Matrix

The equations used to calculate accuracy, precision, and recall are as follows:

1) Accuracy: Accuracy is a metric used to measure how well the model can classify correctly.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN.} \quad (21)$$

2) Precision: Precision is a metric used to measure the proportion of positive predictions that are actually included in the positive category.

$$Precision = \frac{TP}{TP + FP} \quad (22)$$

3) Recall: Recall is a metric used to calculate the proportion of data successfully predicted as positive out of all data that actually belongs to the positive category.

$$Recall = \frac{TP}{TP + FN} \quad (23)$$

4) F1 Score: F1 score is a metric that calculates the harmonic mean between precision and recall.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (24)$$

IV. RESULTS

A. Data Preprocessing

Data preprocessing plays a crucial role in machine learning as it can significantly affect model performance. This process helps to improve the quality and accuracy of the results produced. In this study, the following steps were performed as part of the data preprocessing:

1) Data Selection

The first step in the preprocessing phase involved selecting the relevant data for classification. The dataset was examined, revealing 291 individuals diagnosed with thyroid disease and 3,480 individuals without thyroid disease. This initial selection of data, which included both positive and negative cases, formed the foundation for the subsequent modeling stage, ensuring a balanced and comprehensive dataset for analysis.

2) Data Cleaning

The data cleaning process involved checking for missing values and duplicate data entries. After performing this check, it was found that there were no missing values. However, 63 duplicate records were detected and subsequently removed. After the cleaning process, the dataset contained 3,417 records for the non-diseased class (class 0) and 291 records for the diseased class (class 1). Following this, an outlier analysis was conducted on numerical features such as Age, TSH, TT4, T4U, and FTI, as these features are critical for describing thyroid function. Descriptive statistical analysis was performed to assess the data distribution and detect any extreme values or outliers that could potentially affect the model's results.

3) Data Transformation

To stabilize the feature scales and reduce the influence of outliers, data transformation was carried out. The dataset displayed scale differences and extreme values that could hinder the model's performance, as shown in Table 3.2. To address this, the RobustScaler method was applied, which normalizes the data using the median and interquartile range (IQR) values, ensuring that outliers have less impact. The transformation results, which reflect a more uniform data distribution, are presented in Table 4.

TABLE 4
DATA NORMALIZATION RESULTS

No	Age	Sex	...	TSH	TT4	T4U	FTI
1	-0.419355	0	...	-0.085616	0.583942	0.833333	-0.035714
2	-1	0	...	0.873288	-0.087591	0.027776	0.016773
3	-0.258065	1	...	-0.195205	0.116788	-0.444444	0.357143
...	...	...	...	...	...	...	...
3705	0.645161	0	...	1.215753	0.20438	0.44444	-0.178571
3706	0.580645	1	...	-0.291096	-0.671533	-0.277778	-0.821429
3707	0.322581	0	...	-0.188356	-0.175182	0.444444	-0.642857

Table 4 shows the normalization results using RobustScaler, which stabilizes feature scales and reduces the influence of outliers. This method is based on the median and IQR, enabling it to handle outliers without removing data, making feature scales more uniform.

B. Handling Data Imbalance

After preprocessing, Random Oversampling (ROS) is applied to balance the dataset. In this method, minority class samples are randomly duplicated to increase their number, as calculated by Equation (3). At each iteration, one sample is selected randomly using Equation (4), and the final dataset is formed by combining the majority class, original minority class, and the newly generated ROS samples according to Equation (5). After ROS, the dataset grows from 3,771 to 6,834 samples, achieving a balanced class distribution, as shown in Figure 4.

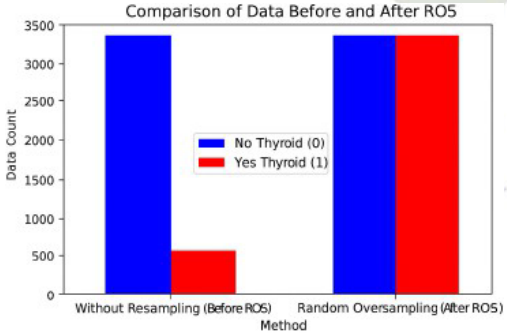

Fig. 4 Implementation of the random oversampling (ROS) method

C. Data Division

After *the* data balancing process, 6,834 data samples were obtained and then divided using the *K-Fold Cross Validation* method. The dataset was divided into 10 folds, with one fold as test data and nine folds as

training data in each iteration, so that all data had the opportunity to become test data alternately. The data distribution in each fold is shown in Table 5

TABLE 5
DATA DIVISION

Fold	Train	Test	y_train (0)	y_train (1)	y_test (0)	y_test (1)
1	6150	684	3057	3093	360	324
2	6150	684	3082	3068	335	349
3	6150	684	3071	3079	346	338
4	6150	684	3067	3083	350	334
5	6151	683	3079	3072	338	345
6	6151	683	3097	3054	320	363
7	6151	683	3096	3055	321	362
8	6151	683	3054	3097	363	320
9	6151	683	3072	3079	345	338
10	6151	683	3078	3073	339	344

With this division, it can be seen that the amount of training and test data is relatively consistent in each fold, and the class distribution between 0 (not affected by thyroid) and 1 (affected by thyroid) remains balanced so that model validation can be carried out fairly on the entire dataset.

D. Backpropagation Neural Network

Backpropagation Neural Network is widely used because it is capable of effectively training neural networks with many layers, and this method has become one of the important foundations in various artificial intelligence applications. Structurally, BPNN consists of three main layers: the input layer (x), the hidden layer (z), and the output layer (y). The architecture of the Backpropagation Neural Network model can be seen in Figure 8.

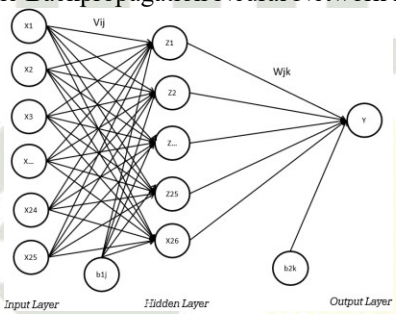

Fig. 5 Backpropagation Neural Network Architecture

Figure 8 shows the BPNN architecture built using 25 attributes from the dataset. The number of neurons in the hidden layer was tested in several variations, namely 26, 38, and 49 neurons. This model uses a sigmoid activation function, with experiments conducted for up to 5,000 and 10,000 epochs. In addition, testing was also conducted using several learning rate values, namely 0.1, 0.01, and 0.001. The BPNN training process was carried out using Equations (6) to (20), while the testing stage applied Equations (6) to (9).

E. Testing and Evaluation

Testing in this study was conducted by trying a number of variations in the BPNN architecture, *learning rate*, and number of epochs. Each combination of parameters was tested to see how these configuration changes affected the model's performance. The results of all testing scenarios, averaged over 3 tests, can be seen in Table 6 below.

TABLE 6
AVERAGE RESULTS OF 3 TESTS

Epoch	K-Fold	Learning Rate	Neuron In Hidden Layer	Accuracy (%)	Precision(%)	Recall (%)	F1 Score (%)
5000	10	0,1	26	99,71%	100%	100%	100%
			38	99,56%	100%	100%	100%
			49	99,71%	100%	100%	100%
		0,01	26	99,71%	100%	100%	100%
			38	99,85%	100%	100%	100%
			49	99.71%	100%	100%	100%
		0,001	26	99,71%	100%	100%	100%
			38	99.41%	100%	100%	100%
			49	99,85%	100%	100%	100%
10000	10	0,1	26	99,56%	100%	100%	100%
			38	99,56%	100%	100%	100%
			49	99,71%	100%	100%	100%
		0,01	26	99,56%	100%	100%	100%
			38	99,71%	100%	100%	100%
			49	99,85%	100%	100%	100%
		0,001	26	99,71%	100%	100%	100%
			38	99,85%	100%	100%	100%
			49	99,56 %	100%	100%	100%

The testing was conducted using three BPNN architecture scenarios, namely [25:26:1], [25:38:1], and [25:49:1], with *learning rate* variations of 0.1, 0.01, and 0.001. Each scenario was also tested using 5,000 and 10,000 epochs.

V. Discussion

Based on the results presented in Table 6, the Backpropagation Neural Network (BPNN) demonstrated remarkable performance, achieving a maximum accuracy of 99.85% under various parameter configurations. This accuracy was consistent across both 5,000 and 10,000 training epochs, indicating that the model maintained stability over extended training periods. Notably, the learning rates of 0.01 and 0.001, along with the hidden layer configurations of 38 and 49 neurons, produced optimal results. The stability observed with smaller learning rates suggests that these values enable the model to learn more gradually, reducing the risk of overshooting the optimal solution and helping the network converge to an optimal state. Thus, fine-tuning the learning rate, number of epochs, and neurons in the hidden layer was critical for achieving the high accuracy observed in this study, as noted by previous research on the significance of parameter tuning in neural network models for improved performance [12]. In a clinical context, this stable optimization behavior is important because it indicates that the model can provide reproducible predictions across different training runs, thereby increasing clinicians' confidence when using it as a decision-support component rather than as a black-box tool with unpredictable variability.

Further examination of the model's performance revealed that the BPNN correctly identified 338 data points from the negative class (class 0) and 344 data points from the positive class (class 1), demonstrating a high degree of classification accuracy. However, there was one instance where a data point from class 0 was misclassified as class 1, which was recorded as a false positive. Importantly, no false negatives were detected, meaning that the model did not mistakenly classify any positive samples as negative. From the perspective of endocrinologists and general practitioners, this pattern is clinically meaningful: the absence of false negatives reduces the risk of undiagnosed thyroid disease, while a small number of false positives would generally lead only to additional laboratory tests or follow-up examinations rather than direct patient harm. The occurrence of a single false positive suggests that while the model performed exceptionally well, there is still a minor possibility for misclassification, though it is relatively rare. Hospital administrators and clinicians could use this information to set clear protocols—for example, confirming model-positive cases with standard laboratory and clinical assessments—to ensure that the impact of occasional false positives is appropriately managed. This indicates the potential for slight improvement in addressing class misclassification in future model iterations, which aligns with the findings of previous studies that highlight the importance of refining model precision [15].

The distribution of correct and incorrect predictions, as shown in the confusion matrix (Figure 6), further supports the model's high level of accuracy. The confusion matrix serves as a valuable tool for understanding how well the model distinguishes between the two classes. With minimal misclassifications, the model was able to effectively learn the underlying patterns in the dataset, which is crucial for tasks such as thyroid disease classification. The fact that the model achieved 99.85% accuracy highlights its potential to assist in clinical settings, where reliable predictions are essential for patient diagnosis and treatment. This is consistent with other studies that have emphasized the importance of high accuracy in medical applications of machine learning [14]. At the same time, stakeholders must remain aware that even highly accurate models can generate occasional misclassifications, so institutional policies should frame the model as an assistive tool that complements, rather than replaces, clinical judgement.

When comparing the performance of this study's model with prior research, the results demonstrate a significant improvement in accuracy. Previous studies, such as the one by [22], reported an accuracy of 90% when using a percentage split and 95% with 5-fold cross-validation. In contrast, the current study achieved an impressive 99.85% accuracy by optimizing key parameters and utilizing 10-fold cross-validation. This improvement emphasizes the importance of parameter tuning in enhancing model performance. By experimenting with various configurations of learning rates, hidden layer neurons, and epochs, this study was able to achieve superior results, underscoring the need for careful optimization in machine learning applications. This approach also demonstrates the potential of cross-validation to enhance model robustness, as confirmed by studies advocating for its use in improving model evaluation [19].

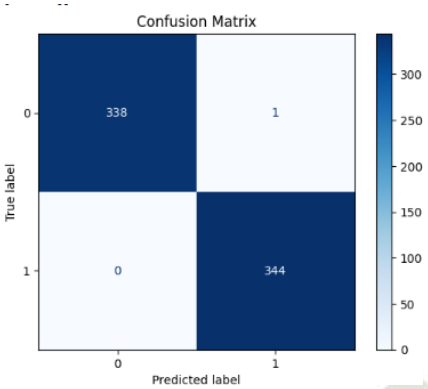

Fig. 6 Confusion Matrix Results s with the Highest Accuracy in Backpropagation Testing

In conclusion, the findings of this study underscore the effectiveness of the Backpropagation Neural Network (BPNN) in classifying thyroid disease with exceptional accuracy. The use of optimized parameters, such as learning rates, the number of epochs, and hidden neurons, contributed significantly to the model's performance. The study also demonstrates the potential of 10-fold cross-validation to provide more reliable and consistent results, compared to other validation methods. Moving forward, further research could explore even more advanced techniques, such as incorporating additional layers or testing with alternative optimization methods like Grid Search, to further enhance the model's classification performance. These directions are supported by previous research that emphasizes continuous improvement and optimization for achieving better classification outcomes in complex datasets [16].

VI. CONCLUSIONS

This study focused on improving thyroid disease classification by applying preprocessing techniques and the Backpropagation Neural Network (BPNN). The dataset used comprised 3,771 patient records, each with 25 clinical and medical attributes. The preprocessing steps included removing duplicates, handling outliers, and normalizing features using RobustScaler. To address the issue of class imbalance, Random Over Sampling (ROS) was employed, increasing the minority class samples from 3,771 to 6,834, thus creating a more balanced dataset. BPNN training was performed with different configurations of hidden neurons (38 and 49), learning rates (0.01 and 0.001), and epochs (5,000 and 10,000), with model performance evaluated through 10-fold cross-validation. The best configuration of the proposed model achieved an accuracy of 99.85%, validating the effectiveness of the combined preprocessing pipeline and parameter tuning in optimizing classification performance for thyroid disease data. These results indicate that the proposed approach is highly suitable to be developed further as an automated decision-support component to assist clinicians in screening and stratifying patients at risk of thyroid disorders.

The use of RobustScaler and ROS contributed significantly to improving model performance by minimizing bias caused by uneven feature scales and class imbalance. The confusion matrix results demonstrated that most samples were correctly classified, with only minimal errors. This pattern, characterized by the absence of false negatives and only a very small number of false positives, is particularly beneficial in clinical practice because it reduces the likelihood of missed thyroid disease cases while keeping the additional follow-up burden at a manageable level. For future work, further improvements can be explored by experimenting with additional hidden layers, varying the number of neurons, testing alternative weight initializations, and employing parameter optimization methods such as Grid Search, Random Search, or Genetic Algorithms. Subsequent studies should also validate the model on multi-center hospital datasets, incorporate additional clinical and socio-demographic variables, and integrate the classifier into hospital information or national health database systems to assess scalability, interoperability, and real-world impact on thyroid disease detection workflows.

REFERENCES

[1] S. Kalra, S. Aggarwal, dan D. Khandelwal, "Thyroid dysfunction and dysmetabolic syndrome: The need for enhanced thyrovigilance strategies," *Int. J. Endocrinol.*, vol. 2021, 2021, doi: 10.1155/2021/9641846.

[2] X. Li, T. Zhang, H. Zhang, S. Liu, dan L. Tian, "Effects of levothyroxine therapy on bone and mineral metabolism in hypothyroidism: a systematic review and meta-analysis," *BMC Endocr. Disord.*, vol. 25, no. 1, 2025, doi: 10.1186/s12902-024-01819-7.

[3] Rifka Aulia Astuti dan Farida Noor Irfani, "Pemeriksaan Imunologi Terhadap Kadar Hormon Thyroid Stimulating Hormone (TSH) Pada Pasien Gangguan Tiroid Di RSUD Panembahan Senopati Periode 2020-2022," *J. Ilm. Kedokt. dan Kesehat.*, vol. 3, no. 1, hal. 280–292, 2024, doi: 10.55606/klinik.v3i1.2631.

[4] N. Komang dan M. Putri, "Analisis Pemeriksaan Laboratorium pada Penyakit Tiroid," *Action Research Literate*, vol. 8, no. 8, Aug. 2024, doi: https://doi.org/10.46799/arl.v8i8.482.

[5] N. K. M. Suryantini, L. L. Putri, B. H. Salim, A. Mawaddah, dan E. Triani, "Gangguan Hormon Tiroid: Hipotiroidisme," *J. Ilmu Kedokt. dan Kesehat.*, vol. 11, no. 6, hal. 1227–1234, 2024, doi: 10.33024/jikk.v11i6.14697.

[6] A. H. Gupta; dan Alice Callahan, "How to Know if You're Iron-Deficient, and What to Do About It ," *Nytimes.com*, Oct. 17, 2023. https://www.nytimes.com/2023/10/17/well/live/iron-deficient-testing-treatment.html

[7] I. P. Hati, "Mewaspadai Komplikasi Akibat Gangguan Tiroid," *Bisnis.com, 2022. https://koran.bisnis.com/read/20241130/462/1820303/penyakit-kelenjar-mewaspadai-komplikasi-akibat-gangguan-tiroid (accessed Dec. 09, 2025).* [8] S. L. Anwar, R. Cahyono, S. Suwardjo, dan H. Hardiyanto, "Extrathyroidal extension and cervical node infiltration are associated with recurrences and shorter recurrence-free survival in differentiated thyroid cancer: a cohort study," *Thyroid Res.*, vol. 15, no. 1, hal. 1–8, 2022, doi: 10.1186/s13044-022-00131-7.

[9] M. S. Salsabilah *et al.*, "Hipotiroidsme : Etiologi, Faktor Risiko Dan Tatalaksana Komprehensif," *J. Kesehat. Tambusai*, vol. 5, no. 4, hal. 13211–13218, 2024, doi: 10.31004/jkt.v5i4.36775.

[10] M. Amalia, C. A. Martafari, dan E. Emiralda, "Faktor Risiko Kejadian Hipertiroid di Rumah Sakit Meuraxa Banda Aceh," *J. Kesehat. Tambusai*, vol. 4, no. 3, hal. 2679–2686, 2023, doi: 10.31004/jkt.v4i3.17490.

[11] M. J. Thomsen, S. L. Andersen, N. M. U. Torp, S. Andersen, dan A. Carlé, "Challenges in subtyping disease entities of thyroid dysfunction following combined immune checkpoint inhibitor therapy: a case report," *Thyroid Res.*, vol. 18, no. 1, hal. 4–9, 2025, doi: 10.1186/s13044-025-00254-7.

[12] J. A. Orozco Torres, A. Medina Santiago, J. M. Villegas Izaguirre, M. Amador García, dan A. Delgado Hernández, "Hypertension Diagnosis with Backpropagation Neural Networks for Sustainability in Public Health," *Sensors*, vol. 22, no. 14, hal. 1–20, 2022, doi: 10.3390/s22145272.

[13] F. Muhajir, F. N. Irfani, and W. S. Hadi, "Gambaran Hasil Pemeriksaan Laboratorium Pada Penderita Gangguan Tiroid Di RSUD Kabupaten Temanggung," *Borneo Journal of Medical Laboratory Technology*, vol. 8, no. 1, pp. 966–973, 2025, doi: https://doi.org/10.33084/bjmlt.v8i1.10809.

[14] A. Raza, F. Eid, E. C. Montero, I. D. Noya, dan I. Ashraf, "Enhanced interpretable thyroid disease diagnosis by leveraging synthetic oversampling and machine learning models," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, 2024, doi: 10.1186/s12911-024-02780-0.

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

[16] M. Salmi, D. Atif, D. Oliva, A. Abraham, dan S. Ventura, *Handling imbalanced medical datasets: review of a decade of research*, vol. 57, no. 10. Springer Netherlands, 2024. doi: 10.1007/s10462-024-10884-2.

[17] T. Tanti, "Random Oversampling, Chi-Square, dan AdaBoost dalam Penanganan Ketidakseimbangan Kelas pada Klasifikasi C5.0," *J. Media Inform. Budidarma*, vol. 7, no. 2, hal. 714, 2023, doi: 10.30865/mib.v7i2.5862.

[18] T. Wongvorachan, S. He, dan O. Bulut, "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining," *Inf.*, vol. 14, no. 1, 2023, doi: 10.3390/info14010054.

[19] L. B. V. de Amorim, G. D. C. Cavalcanti, dan R. M. O. Cruz, "The choice of scaling technique matters for classification performance," *Appl. Soft Comput.*, vol. 133, hal. 1–37, 2023, doi: 10.1016/j.asoc.2022.109924.

[20] A. C. Amutha C dan P. P. P. Periyasamy, "A Weighted Fuzzy-Based Model for Thyroid Prediction in Distributed Environment," *J. Neonatal Surg.*, vol. 14, no. 23S, hal. 377–391, 2025, doi: 10.63682/jns.v14i23s.5757.

[21] P. R. Sihombing, S. Suryadiningrat, D. A. Sunarjo, dan Y. P. A. C. Yuda, "Identifikasi Data Outlier (Pencilan) dan Kenormalan Data Pada Data Univariat serta Alternatif Penyelesaiannya," *J. Ekon. Dan Stat. Indones.*, vol. 2, no. 3, hal. 307–316, 2023, doi: 10.11594/jesi.02.03.07.

[22] N. Henisaniyya *et al.*, "Sistemasi: Jurnal Sistem Informasi Klasifikasi Class Thyroid menggunakan Algoritma ID3 dan Artificial Neural Network (ANN) Classification of Thyroid Class using ID3 Algorithm and Artificial Neural Network (ANN)," vol. 14, no. 1, hal. 2540–9719, 2025, [Daring]. Tersedia pada: http://sistemasi.ftik.unisi.ac.id

[23] U. Khasanah, "Klasifikasi Multi Class Menggunakan Metode Backpropagation Neural Network Untuk Prediksi Stunting Pada Balita," *MathVision J. Mat.*, vol. 7, no. 1, hal. 13–21, 2025, doi: 10.55719/mv.v7i1.1650.

[24] I Wayan Supriana, Gede, Gede, and I Ketut Adian Jayaditya, "Implementasi Artificial Neural Network Menggunakan Backpropagation Pada Prediksi Indeks Pembangunan Manusia," *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*, vol. 13, no. 4, pp. 875–882, 2025, doi: https://doi.org/10.24843/JLK.2025.v13.i04.p14.

[25] M. R. Al Fathan, M. F. Arfa, dan ..., "Algorithm Decission Tree C4. 5 and Backpropagation Neural Network for Smarthpone Price Classification," *... J. Artif. ...*, vol. 5, no. 1, hal. 120–129, 2022, [Daring]. Tersedia pada: http://ejournal.uin-suska.ac.id/index.php/IJAIDM/article/view/19064

[26] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African J. Biomed. Res.*, no. December, hal. 4023–4031, 2024, doi: 10.53555/ajbr.v27i4s.4345.

[27] W. E. S. Ningsih dan E. Ismawatie, "Korelasi Kadar Tsh Dan Kadar Ft4 Pada Pasien Tiroid Di Kabupaten

Bangkalan Tahun 2023," *Plenary Health :* Jurnal Kesehatan Paripurna, vol. 1, no. 3, pp. 244–249, 2023, doi: https://doi.org/10.37985/plenaryhealth.v1i3.592.
[28] W. Nugraha, R. Sabaruddin, dan S. Murni, "Teknik Scaling Menggunakan Robust Scaler Untuk Mengatasi Outlier Data Pada Model Prediksi Serangan Jantung," *Techno.Com*, vol. 23, no. 2, hal. 319–327, 2024, doi: 10.62411/tc.v23i2.10463.
[29] N. C. Nugraha, H. Hikmayanti, J. Indra, dan A. R. Juwita, "Implementasi Metode Resampling Dalam Menangani Data Imbalance Pada Klasifikasi Multiclass Penyakit Thyroid," *Build. Informatics, Technol. Sci.*, vol. 6, no. 2, hal. 890–900, 2024, doi: 10.47065/bits.v6i2.5652.
[30] rahayu deny danar dan alvi furwanti Alwie, A. B. Prasetio, R. Andespa, P. N. Lhokseumawe, dan K. Pengantar, "Tugas Akhir Tugas Akhir," *J. Ekon. Vol. 18, Nomor 1 Maret201*, vol. 2, no. 1, hal. 41–49, 2020.
[31] M. Azhima, I. Afrianty, E. Budianita, dan S. Kurnia Gusti, "Penerapan Metode Backpropagation Neural Network untuk Klasifikasi Penyakit Stroke," *Media Online)*, vol. 4, no. 6, hal. 3013–3021, 2024, doi: 10.30865/klik.v4i6.1956.
[32] H. Hamdani, Z. Arifin, dan A. Septiarini, "Expert System of Dengue Disease Using Artificial Neural Network Classifier," *JUITA J. Inform.*, vol. 10, no. 1, hal. 59, 2022, doi: 10.30595/juita.v10i1.12476.
[33] M. Anshori, M. S. Haris, dan W. Teja Kusuma, "Penerapan Backpropagation Neural Network (BPNN) Untuk Prediksi Kecanduan Smartphone Pada Remaja," *Cices*, vol. 9, no. 2, hal. 192–202, 2023, doi: 10.33050/cices.v9i2.2701.
[34] U. H. Mubarok dan Suprapto, "Perbandingan Akurasi antara Jaringan Saraf Tiruan Backpropagation dan Case-Based Reasoning dalam Diagnosis Penyakit Mata," *Berk. MIPA*, vol. 25, no. 2, hal. 154–164, 2018.
[35] W. Wijiyanto, A. I. Pradana, S. Sopingi, dan V. Atina, "Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa," *J. Algoritm.*, vol. 21, no. 1, hal. 239–248, 2024, doi: 10.33364/algoritma/v.21-1.1618.
[36] H. Hafid, "Penerapan K-Fold Cross Validation untuk Menganalisis Kinerja Algoritma K-Nearest Neighbor pada Data Kasus Covid-19 di Indonesia," *J. Math.*, vol. 6, no. 2, hal. 161–168, 2023, [Daring]. Tersedia pada: http://www.ojs.unm.ac.id/jmathcos
[37] H. Mustafidah dan S. Suwarsito, "Performance of Levenberg-Marquardt Algorithm in Backpropagation Network Based on the Number of Neurons in Hidden Layers and Learning Rate," *JUITA J. Inform.*, vol. 8, no. 1, hal. 29, 2020, doi: 10.30595/juita.v8i1.7150.
[38] Bagus Arya Atmaja, Gery Samuel Gultom, Jhon Hansen Manurung, dan Victor Asido Elyakim P, "Penerapan Metode Backpropagation dalam Jaringan Saraf Tiruan untuk Meningkatkan Prediksi Produksi Kentang di Sumatera," *INSOLOGI J. Sains dan Teknol.*, vol. 3, no. 6, hal. 674–683, 2024, doi: 10.55123/insologi.v3i6.4658.
[39] Y. Muamar dan A. Muhajirin, "Penerapan Jaringan Saraf Tiruan Dengan Metode Backpropagation Untuk Memprediksi Tingkat Kelulusan Mahasiswa Perguruan Tinggi," *Digit. Transform. Technol.*, vol. 4, no. 1, hal. 214–224, 2024, doi: 10.47709/digitech.v4i1.3810.
[40] A. Santriawan, Gunadi Widi Nurcahyo, dan Billy Hendrik, "Prediksi Penjualan Sepeda Motor Yamaha dengan Jaringan Syaraf Tiruan dan Backpropagation (Studi Kasus: CV Sinar Mas)," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 5, no. 1, hal. 185–194, 2024, doi: 10.37859/coscitech.v5i1.6709.
[41] S.-Z. Zhang, S. Chen, and H. Jiang, "A back propagation neural network model for accurately predicting the removal efficiency of ammonia nitrogen in wastewater treatment plants using different biological processes," *Water Research*, vol. 222, p. 118908, Aug. 2022, doi: https://doi.org/10.1016/j.watres.2022.118

KOMUNITAS DOSEN INDONESIA

Editor Decision
KDI/993/LoA-Invoice/2025
10/12/2025

Dears Mr/Mrs/Ms. **Ummy Agustina Putri, Iis Afrianty, Elvia Budianita, Fadhilah Syafria**
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)**, **Application of Backpropagation Neural Network Using Random Oversampling and Robust Scaler for Classification Thyroid** 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

Jl. Flamboyan 2 Blok B3 No. 26 Griya Sangiang Mas - Tangerang 15132 Telp. 081807834703
jurnal.kdi.or.id | kdieco.buss@gmail.com