

**KLASIFIKASI PENYAKIT TIROID
MENGUNAKAN ALGORITMA
*ARTIFICIAL NEURAL
NETWORK***

**OLEH
YAYAN NASIBU
T3121033**

**SKRIPSI
Untuk Memenuhi Salah Satu Syarat Ujian
Guna Memperoleh Gelar Sarjana**



**PROGRAM SARJANA
TEKNIK INFORMATIKA
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2025**

PERSETUJUAN SKRIPSI

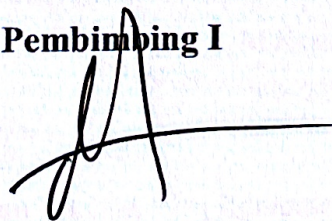
**KLASIFIKASI PENYAKIT TIROID
MENGUNAKAN ALGORITMA
*ARTIFICIAL NEURAL
NETWORK***

**OLEH
YAYAN NASIBU
T3121033**

SKRIPSI

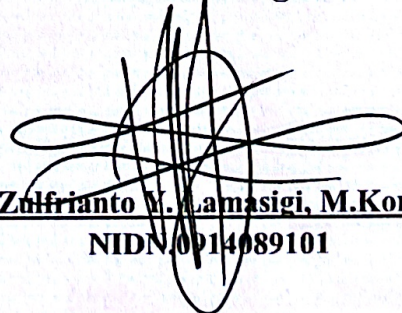
Untuk Memenuhi Salah Satu Syarat Ujian
Guna Memperoleh Gelar Sarjana
Program Studi Teknik Informatika,
Ini Telah Disetujui Oleh Pembimbing

Pembimbing I



**Sudirman S. Panna, M.Kom
NIDN.0924038205**

Pembimbing II



**Zulfrianto W. Lamasigi, M.Kom
NIDN.0914089101**

PENGESAHAN SKRIPSI
KLASIFIKASI PENYAKIT TIROID
MENGGUNAKAN ALGORITMA
ARTIFICIAL NEURAL
NETWORK

Oleh :

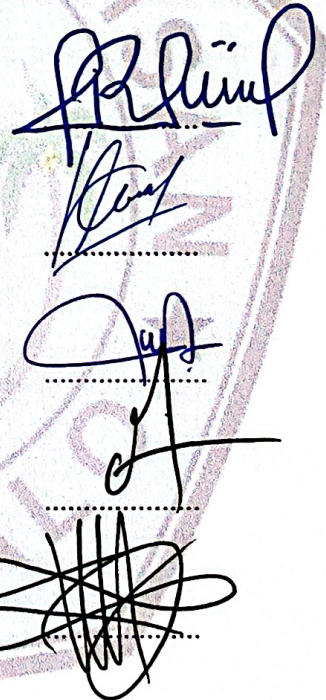
YAYAN NASIBU

T3121033

Diperiksa oleh Panitia Ujian Strata Satu (S1)

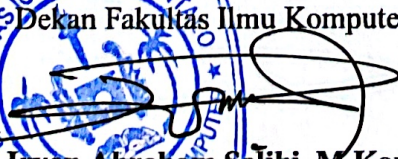
Universitas Ichsan Gorontalo

1. Ketua Penguji
Rezqiwati Ishak, M.Kom
2. Anggota Penguji
Husdi, M.Kom
3. Anggota Penguji
Ir. Misrawaty A. Puspa, M.Kom
4. Pembimbing Utama
Sudirman S. Panna, M.Kom
5. Pembimbing Pendamping
Zulfrianto Y. Lamasigi, M.Kom

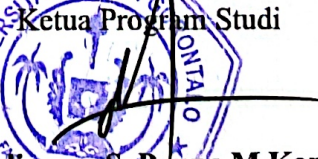


Mengetahui



Dekan Fakultas Ilmu Komputer

Irvan Abraham Salihi, M.Kom
NIDN:0928028101



Ketua Program Studi

Sudirman S. Panna M.Kom
NIDN:0924038205

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa :

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun di perguruan tinggi lainnya.
2. Karya tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan dari Tim Pembimbing.
3. Dalam karya tulis (Skripsi) saya tidak terdapat karya atau pendapat yang telah dipublikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan/sitasi dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma-norma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, Februari 2025

Yang Membuat Pernyataan



Yayan Nasibu

NIM: T3121033

ABSTRACT

YAYAN NASIBU. T3121033. THYROID DISEASE CLASSIFICATION USING ARTIFICIAL NEURAL NETWORK ALGORITHM

Thyroid diseases, such as hypothyroidism and hyperthyroidism, are commonly encountered endocrine disorders that can lead to serious complications if not promptly detected and treated. Conventional diagnostic methods often face difficulties handling complex and non-linear data, thus requiring a more efficient approach. This research aims to develop a thyroid disease classification model using the Artificial Neural Network (ANN) algorithm. The ANN model is constructed using 383 medical records from patients with a history of thyroid disease to train the model in recognizing patterns and classifying thyroid conditions. The evaluation is conducted through Accuracy, Precision, Recall, and F1-Score metrics. The research results indicate that the ANN model obtains a high accuracy of 97%, with significant improvements in accuracy and a reduction in loss as the epochs progressed. Evaluation for both classes (Yes and No) also shows satisfactory results, with high recall and accuracy values. The Confusion Matrix demonstrated the model's ability to classify the entire dataset accurately. This ANN model is effective for classifying thyroid disease, and it has the potential to be applied in clinical decision support systems. However, this research is limited by the restricted size and diversity of the dataset, which could impact the model's generalization capability. For future research, it is recommended to utilize larger and more diverse datasets and to compare performance with other algorithms like Random Forest or Support Vector Machine (SVM) to enhance the accuracy and efficiency of thyroid disease classification models.

Keywords: *thyroid disease, Artificial Neural Network, classification, accuracy.*



ABSTRAK

YAYAN NASIBU. T3121033. KLASIFIKASI PENYAKIT TIROID MENGGUNAKAN ALGORITMA ARTIFICIAL NEURAL NETWORK

Penyakit tiroid, seperti hipotiroidisme dan hipertiroidisme, merupakan gangguan endokrin yang sering ditemukan dan dapat menyebabkan komplikasi serius jika tidak terdeteksi dan ditangani dengan cepat. Metode diagnosis konvensional sering kali kesulitan dalam menangani data yang kompleks dan tidak linier, sehingga dibutuhkan pendekatan yang lebih efisien. Penelitian ini bertujuan untuk mengembangkan model klasifikasi penyakit tiroid menggunakan algoritma Artificial Neural Network (ANN). Model ANN dibangun menggunakan 383 rekam medis pasien dengan riwayat penyakit tiroid untuk melatih model dalam mengenali pola dan mengklasifikasikan kondisi tiroid. Evaluasi dilakukan dengan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model ANN mencapai akurasi tinggi 97%, dengan peningkatan akurasi dan penurunan loss yang signifikan seiring berjalannya epoch. Evaluasi untuk kedua kelas (Ya dan Tidak) juga menunjukkan hasil yang memuaskan, dengan nilai recall dan akurasi yang tinggi. Confusion Matrix membuktikan bahwa model mampu mengklasifikasikan seluruh data dengan benar. Kesimpulannya, model ANN ini efektif dalam klasifikasi penyakit tiroid dan dapat berpotensi diterapkan dalam sistem pendukung keputusan klinis. Namun, penelitian ini memiliki keterbatasan pada jumlah dan keragaman dataset yang terbatas, yang dapat mempengaruhi kemampuan generalisasi model. Untuk penelitian selanjutnya, disarankan untuk menggunakan dataset yang lebih besar dan lebih beragam, serta membandingkan dengan algoritma lain seperti Random Forest atau Support Vector Machine (SVM) untuk meningkatkan akurasi dan efisiensi model dalam klasifikasi penyakit tiroid.

Kata kunci: penyakit tiroid, Artificial Neural Network, klasifikasi, akurasi.



KATA PENGANTAR

Bismillahirrahmanirrahim

Puji syukur penulis panjatkan kehadirat Allah subhanahun wa ta'ala, karena hanya dengan izin dan kuasa-Nya penulis dapat menyelesaikan penyusunan skripsi ini dengan judul: **“Klasifikasi Penyakit Tiroid Menggunakan Algoritma Artificial Neural Network (ANN)”**, sebagai salah satu syarat Ujian Akhir guna memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Penulis menyadari sepenuhnya bahwa skripsi ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, segala keikhlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Bapak Muh Ichsan Gaffar, M.Ak, CTA, ACPA, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan Dan Teknologi (YPIPT) Ichsan Gorontalo.
2. Ibu Dr. Hj. Juriko Abdussamad, M.Si, selaku Rektor Universitas Ichsan Gorontalo.
3. Bapak Irvan Abraham Salihi, S.Kom, M.Kom, selaku Dekan Fakultas Ilmu Komputer.
4. Bapak Sudirman Melangi, M.Kom, selaku Wakil Dekan I Bidang Akademik Fakultas Ilmu Komputer.
5. Ibu Irma Surya Kumala Idris, S.Kom, M.Kom, selaku Wakil Dekan II Bidang Administrasi Umum dan Keuangan Universitas Ichsan Gorontalo.
6. Bapak Sudirman S. Panna, M.Kom, selaku Pembimbing Utama yang telah membimbing penulis selama mengerjakan penyusunan skripsi ini.
7. Bapak Zulfrianto Y. Lamasigi, M.Kom, selaku Pembimbing Pendamping yang telah membimbing penulis selama mengerjakan penyusunan skripsi ini.

8. Bapak dan Ibu dosen yang telah mendidik dan membimbing penulis selama perkuliahan.
9. Superhero dan panutanku Ayahanda dan pintu surgaku Ibunda. Terima kasih atas segala pengorbanan dan tulus kasih yang diberikan, serta senantiasa memberikan yang terbaik, tak kenal lelah mendoakan, memberikan perhatian, dukungan hingga penulis mampu menyelesaikan studinya sampai meraih gelar sarjana. Semoga ayah dan ibu sehat, panjang umur dan bahagia selalu.
10. Kepada seseorang yang selalu bersama penulis dan tidak bisa penulis sebut namanya. Terima kasih untuk semangat dan perhatian yang diberikan saat proses penyusunan skripsi ini. Ternyata dengan adanya anda di kehidupan penulis memberikan cukup motivasi untuk terus maju dan berproses menjadi pribadi yang mengerti apa itu pengalaman, pendewasaan, sabar dan menerima arti perjuangan sebagai bentuk proses penempaan menghadapi dinamika hidup. Terima kasih telah menjadi bagian menyenangkan dari pendewasaan ini.
11. Teman-teman jurusan teknik informatika, khususnya teman-teman kelas reguler pagi B. Terima kasih telah memberi dukungan, solusi, semangat dan motivasi selama ini serta doa yang terbaik untuk penulis.

Semoga Allah SWT, melimpahkan balasan atas jasa-jasa mereka kepada kami. Penulis menyadari sepenuhnya bahwa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif. Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Aamiin.

Gorontalo, Februari 2025

Penulis

DAFTAR ISI

PERSETUJUAN SKRIPSI.....	ii
PENGESAHAN SKRIPSI.....	iii
PERNYATAAN SKRIPSI.....	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	3
BAB II LANDASAN TEORI	5
2.1 Tinjauan Studi	5
2.2 Tinjauan Pustaka.....	7
2.2.1 Penyakit Tiroid.....	7
2.2.2 Data Mining.....	7
2.2.3 Tahapan Data Mining	8
2.2.4 Proses Data Mining	10
2.2.5 Min-Max Normalization	12
2.2.6 Algoritma Artificial Neural Network	12
2.2.7 Arsitektur Jaringan Syaraf Tiruan.....	13

2.2.8 Cara Kerja Jaringan Syaraf Tiruan	16
2.2.9 <i>Confusion Matrix</i>	17
2.2.10 Penerapan Metode	19
2.3 Kerangka Pikir	22
BAB III METODE PENELITIAN.....	22
3.1 Jenis, Metode, Subyek, Objek, Waktu, dan Lokasi Penelitian	23
3.2 Pengumpulan Data	23
3.3 Analisis Data	25
3.3.1 Pra Pengolahan	25
3.4 Pelatihan	25
3.4.2 Validasi data	25
3.4.2 Evaluasi Model	26
BAB IV HASIL PENELITIAN.....	28
4.1 Hasil Pengumpulan Data	28
4.2 Hasil Pengolahan Data	30
4.2.1 Pembagian data <i>training</i> dan data <i>testing</i>	30
4.2.2 Normalisasi	30
4.2.3 Proses dengan ANN	32
4.3 <i>Confusion Matrix</i>	37
BAB V PEMBAHASAN	40
5.1 Pembahasan Pengumpulan Data	40
5.3 <i>Validation Akurasi</i> dan <i>Validation Loss</i>	40
5.4 <i>Confusion Matrix</i>	41
BAB VI PENUTUP	43
6.1 Kesimpulan	43
6.2 Saran	43
DAFTAR PUSTAKA	44

DAFTAR GAMBAR

<i>Gambar 2.1 Tahapan Data Mining</i> [12].....	9
<i>Gambar 2.2 Representasi Neural Network</i> [14].	12
<i>Gambar 2. 3 Fungsi aktivasi idenstitas $f(x) = x$</i> [14].	13
<i>Gambar 2.4 Jaringan lapisan tunggal</i> [15].....	14
<i>Gambar 2.5 Jaringan lapisan jamak</i> [15].	15
<i>Gambar 2.6 Jaringan lapisan kompetitif</i> [15].	15
<i>Gambar 2.7 Cara kerja jaringan syaraf tiruan</i> [15].	16
<i>Gambar 2.8 Kerangka Pikir</i>	22
<i>Gambar 4.1 Arsitektur ANN</i>	32
<i>Gambar 5.1 Hasil Pemodelan</i>	40
<i>Gambar 5.2 Validation Akurasi dan Validation Loss</i>	41
<i>Gambar 5.3 Confusion Matrix</i>	42

DAFTAR TABEL

<i>Tabel 2.1</i> Penelitian Terkait[9].	5
<i>Tabel 2.2</i> <i>Confusion Matrix</i> [16].....	18
<i>Tabel 3.1</i> Variabel Data	24
<i>Tabel 4.1</i> Dataset Awal Pasien.....	28
<i>Tabel 4.2</i> Sampel data.....	31
<i>Tabel 4.3</i> Hasil Normalisasi.....	31
<i>Tabel 4.4</i> Inisialisasi Bobot dan Bias.....	33
<i>Tabel 4.5</i> <i>Testing</i>	37

DAFTAR LAMPIRAN

Lampiran 1: Data Yang Dikumpulkan.....	47
Lampiran 2: Kode Program.....	51
Lampiran 3: Riwayat Hidup Peneliti.....	37

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kelenjar tiroid adalah kelenjar yang terletak di leher. Kelenjar yang berfungsi menghasilkan hormon tiroid untuk mengatur metabolisme tubuh. Gangguan pada kelenjar tiroid dan hormon tiroid dapat menimbulkan gejala penyakit tiroid yang berbeda-beda, tergantung jenis dan penyebabnya[1]. Gejala yang umumnya timbul akibat penyakit tiroid adalah munculnya benjolan di leher[1]. Bahaya penyakit kelenjar tiroid dapat diderita oleh siapapun, termasuk pada anak usia 7-12 tahun. Apabila tidak segera diobati, penyakit ini bisa berdampak fatal yaitu gagal jantung, mata menonjol, kesehatan mental, kemandulan, bahkan sampai koma[2].

World Health Organization (WHO) melaporkan bahwa, dengan total 586.202 kasus, kanker tiroid menempati urutan ke-7 dalam jumlah kasus terbanyak dari semua jenis kanker diseluruh dunia pada tahun 2020. Kasus ini terus meningkat dalam lima tahun terakhir, mencapai 1.984.927 kasus. Dalam lima tahun terakhir, ada 38.650 kasus kanker tiroid di Indonesia, tergantung pada usia dan jenis kelamin. Pada tahun 2020, kanker tiroid menempati urutan ke-12 dari semua jenis kanker, dengan 13.114 kasus dan 2.224 kematian [3].

Berdasarkan penelitian terdahulu, penelitian dengan judul Perbandingan kinerja *Algoritma Naive Bayes* dan KNN yang merupakan salah satu algoritma dengan proses yang tidak terlalu rumit, dan proses yang cepat. Sehingga dapat memecahkan permasalahan dengan menggunakan data penyakit tiroid. Dengan perbandingan sebuah *accuracy* pada sebuah algoritma, *Algoritma Naive Bayes* dengan data testing 63% dan data training 64% sedangkan akurasi yang dihasilkan pada algoritma *K-Nearest Neighbors* (KNN) adalah 63%[4].

Penelitian mengenai Klasifikasi Penderita Kanker Paru Paru Menggunakan Algoritma *Artificial Neural Network* (ANN). Penelitian ini bertujuan mengklasifikasikan penderita kanker paru, dan menghasilkan akurasi training 92.79% dengan presisi 86.98% dan akurasi testing sebesar 95.12% dengan presisi

90.23%[5]. Penelitian dengan klasifikasi penyakit jantung juga pernah dilakukan yang mencakup pengukuran performa (akurasi, presisi, recall dan f-score) metode ANN dengan 304 data pasien penyakit jantung yang diperoleh dari pusat dataset Kaggle. Hasil pengukuran performa diperoleh nilai akurasi 73.77%, presisi 80.43%, *recall* 84.09% dan *f-score* sebesar 82.22%[6].

Penelitian berjudul Penerapan Metode ANN untuk Prediksi Gagal Jantung, penelitian ini bertujuan membuat model *Artificial Intelligence* yang bisa memprediksi penyakit gagal jantung. Menggunakan metode *Artificial Neural Network* yang algoritma yang mencontoh cara kerja pada jaringan saraf biologis di otak manusia. Dari penelitian tersebut berhasil mendapatkan hasil akurasi terbaik yaitu sebesar 97% [7].

Penelitian mengenai Prediksi Penyebaran Kasus Tuberklorosis dengan Metode ANN. Sampel penelitian ini berjumlah 5.148 responden, data training yang digunakan 75% dari data dan 25% untuk data testing. Hasil penelitian tersebut berhasil meemprediksi penyebaran kasus Tuberkulosis di kota makasar dengan tingkat akurasi 97.59% [8].

Berdasarkan penelitian terdahulu diatas, telah membuktikan bahwa metode ANN (*Artificial Neural Network*) dapat menghasilkan akurasi yang tinggi. Oleh karena itu, solusi dalam penelitian ini yaitu menerapkan metode ANN sebagai alat untuk mengklasifikasi. Mengingat bahwa metode ini telah terbukti berhasil dalam berbagai bidang. Berdasarkan pemaparan tersebut peneliti tertarik melakukan penelitian dengan judul “Klasifikasi Penyakit Tiroid Menggunakan Metode *Artificial Neural Network*”

Penelitian ini menerapkan klasifikasi penyakit tiroid berdasarkan kecenderungan kekambuhan setelah pengobatan yang menggunakan data penyakit tiroid dengan total keseluruhan data yang akan digunakan berjumlah 383 dataset, dengan melakukan klasifikasi yang akurat, diharapkan dapat memberikan wawasan yang lebih baik bagi praktisi kesehatan dalam memilih metode pengobatan yang optimal dan meminimalkan risiko kekambuhan. Selain itu, hasil penelitian ini dapat memberikan kontribusi dalam pengembangan strategi monitoring jangka panjang yang lebih baik bagi penderita penyakit tiroid.

Penelitian ini bertujuan untuk membuat model klasifikasi penyakit tiroid yang lebih akurat dan efisien menggunakan algoritma ANN. Hasil penelitian ini diharapkan dapat membantu dalam diagnosis penyakit tiroid dan dapat digunakan dalam sistem pendukung keputusan klinis untuk membantu dokter membuat diagnosis yang lebih cepat dan tepat.

1.2 Identifikasi Masalah

- 2.1 Tingkat akurasi diagnosis penyakit tiroid menggunakan metode konvensional masih kurang memadai, terutama dalam menangani data yang kompleks dan tidak linier. Hal ini dapat menyebabkan keterlambatan dalam deteksi dan penanganan yang tepat bagi pasien.
- 2.1 Implementasi *Artificial Neural Network* (ANN) dalam klasifikasi penyakit tiroid memiliki potensi untuk meningkatkan akurasi dan efisiensi diagnosis penyakit tiroid.

1.3 Rumusan Masalah

1. Bagaimana membangun model klasifikasi penyakit tiroid menggunakan algoritma *Artificial Neural Network* (ANN) yang dapat memberikan tingkat akurasi yang efisien dalam diagnosis?
2. Bagaimana evaluasi kinerja dari algoritma *Artificial Neural Network* pada klasifikasi penyakit tiroid?

1.4 Tujuan Penelitian

1. Mengembangkan model klasifikasi penyakit tiroid menggunakan algoritma *Artificial Neural Network* (ANN) yang dapat memberikan tingkat akurasi yang efisien dalam proses diagnosis.
2. Mengetahui kinerja dari algoritma *Artificial Neural Network* untuk klasifikasi penyakit tiroid.

1.5 Manfaat Penelitian

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan ilmu pengetahuan dalam bidang diagnosis penyakit dengan menggunakan metode machine learning, khususnya *Artificial Neural Network* (ANN). Hasil penelitian ini dapat menambah wawasan mengenai efektivitas dan

efisiensi penggunaan ANN dalam klasifikasi penyakit tiroid. Menghasilkan referensi bagi penelitian selanjutnya yang terkait dengan klasifikasi penyakit menggunakan algoritma machine learning, terutama dalam hal implementasi dan optimasi model ANN untuk data medis.

2. Manfaat Praktis

Memberikan alternatif solusi yang lebih efisien dan akurat bagi praktisi medis dalam mendiagnosis penyakit tiroid, sehingga dapat membantu dalam pengambilan keputusan klinis yang lebih cepat dan tepat. Model yang dihasilkan dari penelitian ini dapat diimplementasikan dalam sistem pendukung keputusan klinis, yang pada akhirnya dapat meningkatkan kualitas layanan kesehatan dengan menyediakan alat bantu diagnosis yang lebih handal untuk penyakit tiroid.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Berikut adalah beberapa penelitian terkait yang telah dilakukan dalam bidang klasifikasi penyakit tiroid menggunakan berbagai metode kecerdasan buatan dan algoritma pembelajaran mesin :

Tabel 2.1 Penelitian Terkait[9].

Peneliti	Judul Penelitian	Hasil
Rahman, M., Chowdhury, A., & Khan, M. (2020)[9].	Klasifikasi penyakit tiroid berdasarkan mesin vektor pendukung.	Penelitian ini menggunakan metode <i>Support Vector Machine</i> (SVM) untuk mengklasifikasikan penyakit tiroid. Masalah yang dihadapi adalah rendahnya akurasi pada dataset kecil. Metode SVM menunjukkan akurasi sebesar 89%. Namun, kinerja SVM masih dipengaruhi oleh pemilihan parameter yang tepat dan rentan terhadap <i>overfitting</i> pada data yang kompleks.
Azizah, N., Sari, R., & Santoso, F. (2021)[9].	Klasifikasi penyakit tiroid menggunakan algoritma <i>k-nearest neighbor</i>	Penelitian ini menggunakan algoritma <i>k-Nearest Neighbor</i> (k-NN) untuk klasifikasi penyakit tiroid. Masalah utama yang dihadapi adalah kinerja yang kurang efisien saat menghadapi dataset besar. Hasil menunjukkan akurasi sebesar 85%, tetapi kinerja k-NN terpengaruh oleh jumlah tetangga yang dipilih serta kurang tangguh terhadap noise pada data.
Gibson, T., Patel, R., & Yadav, S.	Penerapan jaringan saraf tiruan dalam diagnosis penyakit	Penelitian ini menggunakan algoritma <i>Artificial Neural Network</i> (ANN) untuk klasifikasi

Peneliti	Judul Penelitian	Hasil
(2020)[9].	tiroid: Sebuah studi perbandingan	penyakit tiroid. Masalah utama yang diatasi adalah kompleksitas data non-linier dalam diagnosis tiroid. Dengan menggunakan ANN, penelitian ini berhasil mencapai akurasi sebesar 91%, namun membutuhkan waktu pelatihan yang lebih lama dibandingkan metode tradisional.
Setiawan, A., Hartanto, W., & Lestari, D. (2023)[9].	Mengoptimalkan jaringan saraf tiruan untuk diagnosis penyakit tiroid menggunakan kumpulan data besar	Penelitian ini berfokus pada pengembangan model ANN yang dioptimalkan untuk dataset besar. Masalah yang dihadapi adalah efisiensi ANN pada dataset besar. Dengan teknik optimasi seperti <i>backpropagation</i> yang ditingkatkan, penelitian ini mencapai akurasi yang lebih tinggi, yakni 94%, dengan waktu pelatihan yang lebih efisien.
Jamaludin, I., Rahman, S., & Hasan, M. (2022)[9].	Analisis perbandingan <i>machine learning</i> untuk klasifikasi penyakit tiroid.	Penelitian ini membandingkan beberapa algoritma pembelajaran mesin seperti SVM, k-NN, dan ANN untuk klasifikasi penyakit tiroid. Masalah yang dihadapi adalah perbedaan kinerja algoritma untuk dataset yang berbeda. Hasilnya, ANN memiliki akurasi tertinggi (92%), diikuti oleh SVM (88%) dan k-NN (84%). ANN dinilai lebih unggul dalam menghadapi data kompleks dan non-linier.

2.2 Tinjauan Pustaka

2.2.1 Penyakit Tiroid

Penyakit tiroid adalah gangguan pada kelenjar tiroid yang dapat menyebabkan perubahan signifikan dalam metabolisme tubuh. Ada beberapa jenis penyakit tiroid yang umum, termasuk[10]:

- *Hipotiroidisme*: Kondisi ketika kelenjar tiroid tidak memproduksi hormon tiroid yang cukup. Gejalanya meliputi kelelahan, penambahan berat badan, dan depresi[10].
- *Hipertiroidisme*: Kondisi ketika kelenjar tiroid memproduksi hormon tiroid dalam jumlah yang berlebihan. Gejalanya meliputi penurunan berat badan yang tidak diinginkan, detak jantung cepat, dan kegelisahan[10].

Penanganan dan diagnosis penyakit tiroid memerlukan pemeriksaan laboratorium dan analisis data pasien. Dalam beberapa dekade terakhir, teknologi kecerdasan buatan telah digunakan untuk membantu proses diagnosis penyakit ini dengan klasifikasi otomatis berdasarkan data medis seperti tingkat hormon tiroid[10].

2.2.2 Data Mining

Data Mining adalah proses ekstraksi pengetahuan yang bermanfaat dan berharga dari dataset yang besar, kompleks, dan tidak terstruktur. Tujuan utama dari data mining adalah untuk menemukan pola, hubungan, atau informasi yang tersembunyi dalam data yang dapat digunakan untuk pengambilan keputusan yang lebih baik. Teknik data mining mencakup berbagai metode analisis statistik, pembelajaran mesin, dan kecerdasan buatan untuk menggali pengetahuan yang terkandung dalam data. Data mining memiliki peran krusial dalam berbagai industri dan bidang, termasuk bisnis, ilmu pengetahuan, kesehatan, keuangan, dan lainnya, untuk mendukung pengambilan keputusan yang lebih tepat, efisien, dan efektif[11].

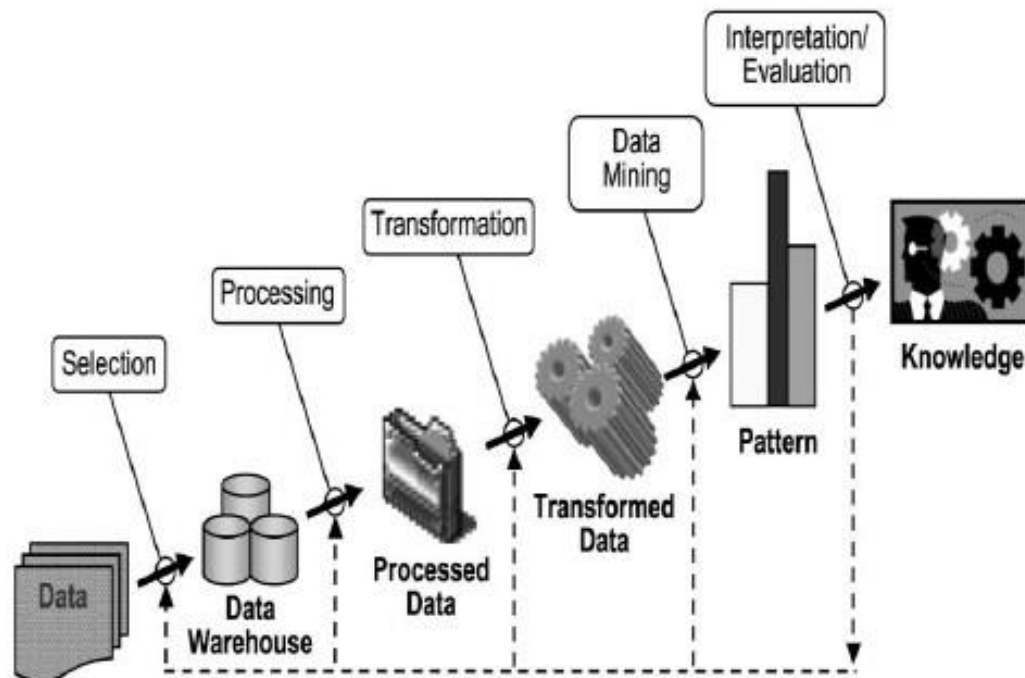
Pemilihan dan persiapan data, pemodelan, evaluasi, dan interpretasi hasil adalah beberapa tahap dalam proses data mining. Tahap pertama adalah pemahaman data, di mana kumpulan data dipelajari secara menyeluruh untuk memahami struktur, fitur, dan kualitasnya. Tahap sebelumnya dari pemrosesan data adalah pembersihan, integrasi, dan transformasi data ke dalam format yang

sesuai untuk analisis. Kemudian dibangun model data mining. dengan menggunakan metode seperti klasifikasi, klasterisasi, regresi, dan asosiasi, sesuai dengan tujuan analisis dan jenis data yang ada saat ini. Setelah pembuatan model, dilakukan evaluasi untuk mengevaluasi validitas dan kinerja model. Terakhir, hasil Data mining digunakan untuk membuat keputusan[11].

Prediksi perilaku konsumen, analisis risiko kredit, deteksi kecurangan, segmentasi pasar, pengoptimalan rantai pasokan, dan banyak lagi adalah beberapa contoh aplikasi data mining. Karena Dengan kemajuan teknologi dan akumulasi data yang terus meningkat, data sangat penting Mining menjadi lebih baik dalam menemukan peluang dan memecahkan masalah, dan meningkatkan kinerja bisnis dan efisiensi pengambilan keputusan di berbagai bidang[11].

2.2.3 Tahapan Data Mining

Proses data mining terdiri dari beberapa tahap. Ini dimulai dengan pemilihan data dari berbagai sumber ke data target, tahap preprocessing untuk meningkatkan kualitas data, transformasi, dan tahap interpretasi dan evaluasi. Semua tahap ini menghasilkan pengetahuan baru yang diharapkan dapat memberikan kontribusi yang lebih baik. Menurut Fayyad (1996), ini dijelaskan secara rinci[12]:



Gambar 2.1 Tahapan Data Mining[12].

1. Data selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional[12].

2. Pre-processing / cleaning

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data[12].

3. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data[12].

4. *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan[12].

5. *Interpretation / evaluation*

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya[12].

2.2.4 Proses Data Mining

Proses yang umumnya dilakukan pada data mining yaitu : deskripsi, prediksi, estimasi, klasifikasi, clustering, dan asosiasi. Secara jelas data mining dijelaskan sebagai berikut[12] :

1. Deskripsi

Tujuan dari deskripsi adalah untuk menemukan pola data yang berulang dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat dimengerti oleh para ahli dalam domain aplikasinya. Agar dengan efektif meningkatkan tingkat pengetahuan sistem, aturan harus mudah dipahami. Tugas deskriptif adalah tugas yang sering dibutuhkan dalam teknik *postprocessing* untuk melakukan validasi dan menjelaskan hasil proses data mining. *Postprocessing* adalah proses yang digunakan untuk memastikan hanya hasil yang valid dan bermanfaat yang dapat digunakan oleh pihak yang berkepentingan[12].

2. Prediksi

Prediksi dan klasifikasi mirip. Namun, data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa depan. Tugas prediksi seperti memprediksi pengurangan pelanggan dalam waktu dekat dan memprediksi harga saham dalam tiga bulan yang akan datang[12].

3. Estimasi

Prediksi dan estimasi hampir sama, kecuali variabel target estimasi lebih ke arah numerik daripada ke arah kategori. Model dibangun dengan menggunakan daftar lengkap yang menunjukkan nilai variabel target sebagai nilai prediksi. Nilai variabel prediksi kemudian digunakan untuk mengestimasi nilai variabel target pada peninjauan berikutnya. Sebagai contoh, akan dihitung tekanan darah sistolik pada pasien rumah sakit berdasarkan umur, jenis kelamin, berat badan, dan level sodium darah. Nilai variabel prediksi dan tekanan darah sistolik akan dihubungkan untuk membuat model estimasi[12].

4. Klasifikasi

Klasifikasi adalah proses menemukan model atau fungsi yang mendeskripsikan data dan membedakan data ke dalam kelas-kelas. Proses ini melibatkan pemeriksaan karakteristik objek untuk memastikan bahwa objek tersebut dimasukkan ke dalam salah satu kelas yang sudah didefinisikan sebelumnya[12].

5. *Clustering*

Pengelompokan data ke dalam kelas objek yang sama tanpa bergantung pada kelas data tertentu dikenal sebagai clustering. Sebuah kluster adalah kumpulan daftar yang memiliki karakteristik yang sama dan tidak memiliki karakteristik yang sama dengan daftar dalam kluster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang sebanding. Semakin banyak kemiripan objek dalam suatu cluster dan lebih banyak perbedaan yang ada di antara masing-masing cluster, semakin baik kualitas analisis cluster[12].

6. Asosiasi

Dalam data mining, tugas asosiasi adalah menemukan fitur yang muncul dalam suatu waktu. Analisis keranjang belanja, atau analisis pasar keranjang, adalah istilah yang lebih umum di dunia bisnis. Salah satu tujuan tugas asosiasi adalah untuk menemukan aturan yang dapat digunakan untuk mengukur hubungan antara dua atau lebih fitur[12].

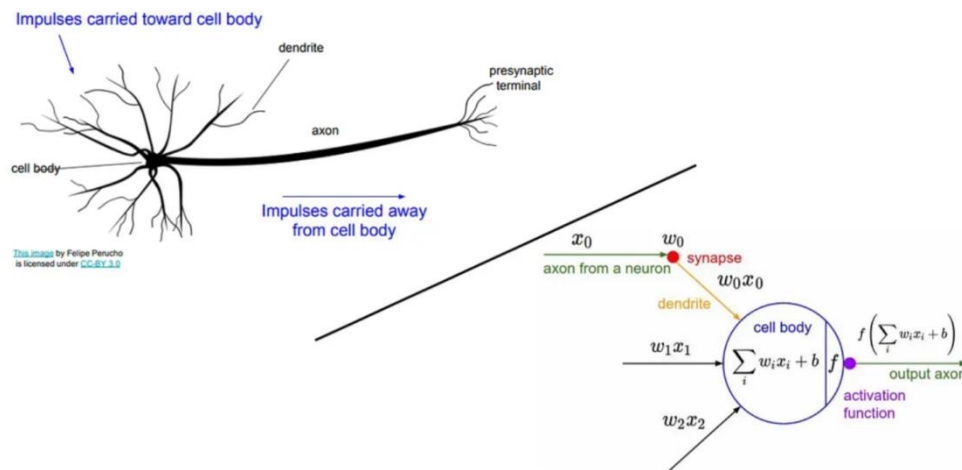
2.2.5 Min-Max Normalization

Terdapat *outlier* dan anomali pada data bisa berdampak buruk pada *Noisy Data* (ND) dan hasilnya intinya, *outlier* mestinya dibuang dengan cara mendeteksinya lebih dahulu misalnya dengan model klasifikasi. Namun cara lain yang klasik dan dapat diterapkan untuk mengurangi *Noisy* pada data yaitu dengan melakukan normalisasi pada data tersebut dengan menggunakan pendekatan *Min-Max Normalization*[13].

$$X_{ij} = \frac{(X_{ij} - X_{minj})}{(X_{maxj} - X_{minj})} (New_{maxj} - New_{minj}) + New_{minj}$$

2.2.6 Algoritma Artificial Neural Network

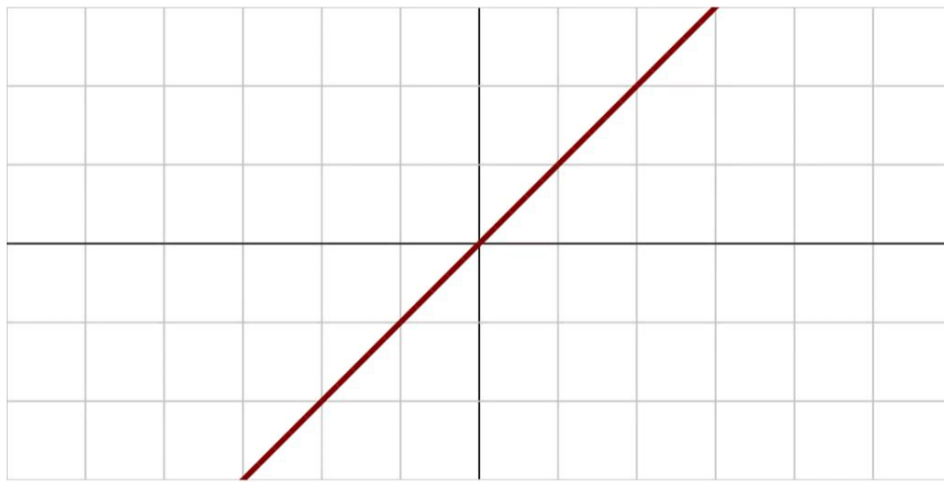
Jaringan Saraf Tiruan (JST) atau *Artificial Neural Network* (ANN) adalah kelompok algoritma yang mirip dengan jaringan saraf otak manusia, di mana neuron terhubung satu sama lain untuk memproses informasi. Tujuan utama ANN adalah untuk membuat komputer dapat berpikir seperti otak manusia, memecahkan masalah, dan belajar[14].



Gambar 2.2 Representasi Neural Network[14].

Pada Gambar 2.1, neuron menerima tiga input (x_0 , x_1 , x_2), masing-masing dikalikan dengan variabel yang dikenal sebagai "berat" (w_0 , w_1 , w_2), dan kemudian ketiga input tersebut digabungkan. Setiap koneksi neuron memiliki "bobot" tertentu, dan manfaatnya akan berubah selama proses pembelajaran sampai model ANN yang dibuat mendekati *output* yang diinginkan[14].

Setelah itu, bias b dapat dimasukkan ke dalam jumlah sebelumnya. Nilai bias ini tidak berasal dari lapisan input; sebaliknya, bias yang serupa dengan intersep dalam persamaan linear ditambahkan untuk meningkatkan akurasi hasil perhitungan. Setelah semua perhitungan dilakukan, neuron akan dimasukkan ke dalam fungsi aktivasi. Fungsi ini menentukan apakah neuron harus aktif atau tidak[14].



Gambar 2. 3 Fungsi aktivasi idenstitas $f(x) = x$ [14].

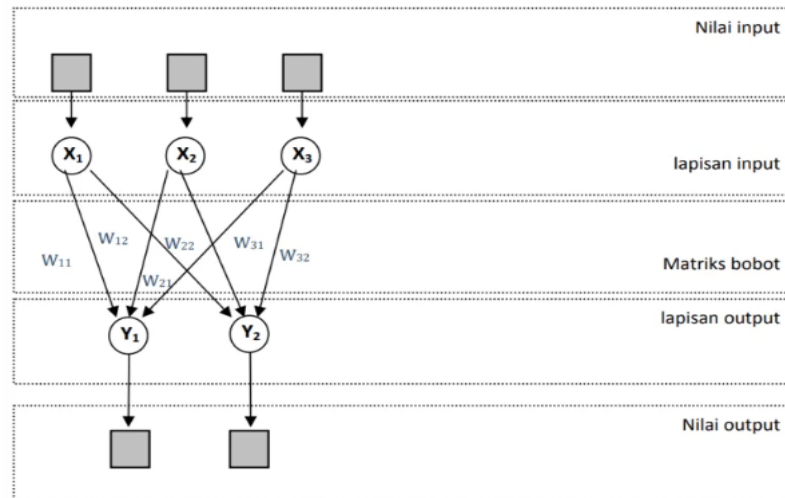
Semua neuron dalam ANN terhubung satu sama lain untuk menghasilkan *output* dari input yang diberikan[14].

2.2.7 Arsitektur Jaringan Syaraf Tiruan

Beberapa arsitektur jaringan yang sering dipakai dalam JST antara lain :

1. Jaringan Lapisan Tunggal (*Single Layer Network*)

Jaringan tunggal hanya memiliki satu lapisan bobot-bobot yang terhubung. Jika input kemudain diterima, jaringan ini dapat mengelolahnya menjadi *output* tanpa melewati lapisan tersembunyi[15].

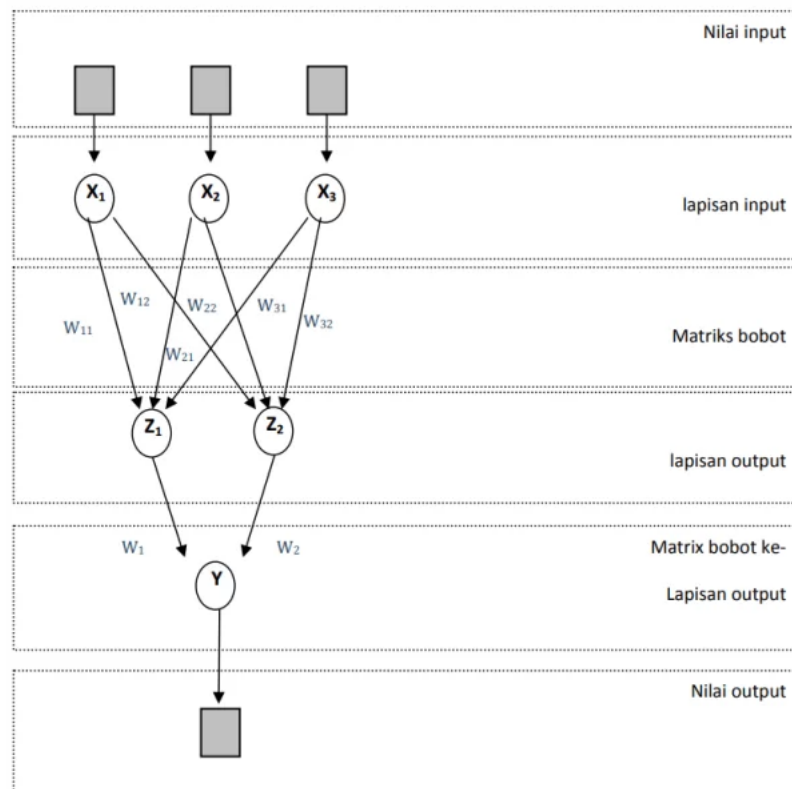


Gambar 2.4 Jaringan lapisan tunggal[15].

Gambar di atas menunjukkan jaringan syaraf dengan tiga neuron (X_1 , X_2 , dan X_3) di lapisan input dan dua neuron (Y_1 , Y_2) di lapisan keluar. Kekuatan hubungan antara neuron di kedua lapisan ditentukan oleh bobot yang disesuaikan[15].

2. Jaringan Lapisan Jamak (*Multi Layer Network*)

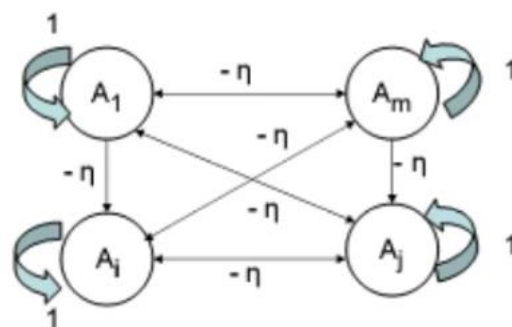
Jaringan lapisan jamak adalah peningkatan dari jaringan lapisan tunggal. Jaringan ini memiliki unit input dan *output* serta unit tambahan yang dikenal sebagai lapisan tersembunyi atau lapisan tersembunyi. Dengan pembelajaran yang lebih kompleks, jaringan banyak lapisan ini dapat menyelesaikan masalah yang lebih sulit daripada jaringan banyak lapisan tunggal[15].



Gambar 2.5 Jaringan lapisan jamak[15].

3. Jaringan Lapisan Kompetitif (*Competitive Layer Net*)

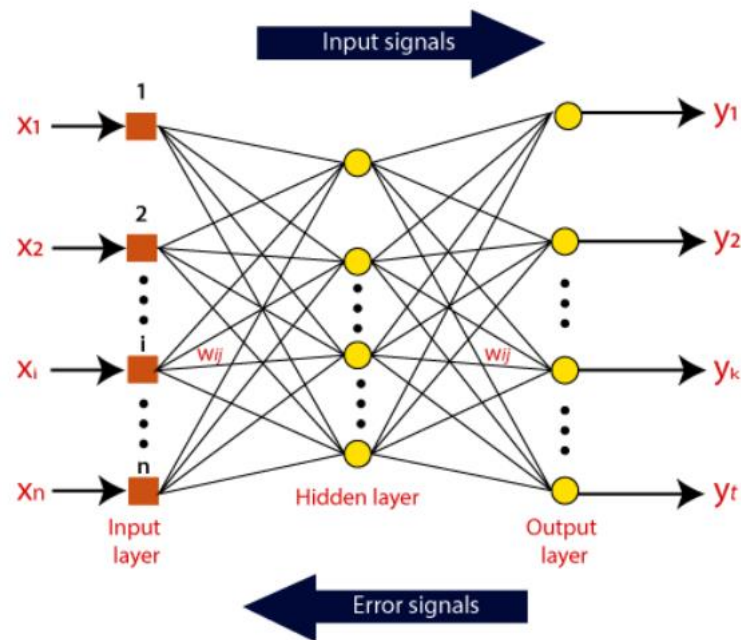
Jaringan lapisan kompetitif memiliki karakteristik yang sebanding dengan jaringan lapisan tunggal atau ganda. Namun, ada neuron *output* (juga dikenal sebagai *loop feedback*) yang memberikan sinyal pada unit input[15].



Gambar 2.6 Jaringan lapisan kompetitif[15].

2.2.8 Cara Kerja Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan mendapatkan sinyal dari sumber luar. Sumber luar dapat berupa pola atau gambar yang diwakili dalam vektor. Selanjutnya, untuk setiap n jumlah input, notasi $X(n)$ digunakan untuk menentukan input secara matematis[15].



Gambar 2.7 Cara kerja jaringan syaraf tiruan[15].

Selanjutnya, masing-masing masukan dikalikan dengan bobot yang disesuaikan. Dalam JST, semua input berbobot diringkas dalam unit koputasi, dan bobot ini biasanya merupakan representasi kekuatan interkoneksi neuron. Bias ditambahkan ketika jumlah bobot sama dengan 0, untuk membuat *output* tidak nol atau sesuatu yang lain untuk meningkatkan reaksi sistem. Bias memiliki input yang sama dan nilai bobotnya sama dengan 1[15].

Nilai maksimum tertentu di banchmark dan total input berbobot dilewatkan oleh fungsi aktivasi untuk menjaga respons dalam batas nilai yang diinginkan. Nilai total input berbobot dapat berada dalam kisaran 0 hingga tak terhingga positif $(+\infty)$ [15].

Fungsi aktivasi non-linier menentukan aktivasi neural tiruan. Suatu nilai ambang, atau batas, digunakan untuk menunjukkan tingkat aktivasi. Fungsi

tangga, lereng, dan sigmoid adalah fungsi aktivasi yang paling umum digunakan[15].

2.2.9 Confusion Matrix

Confusion matrix, juga dikenal sebagai matriks kebingungan, adalah alat yang digunakan untuk menunjukkan kinerja model klasifikasi pada data uji dengan hasil yang sudah diketahui sebelumnya. Kita dapat dengan mudah membayangkan sebuah permainan di mana dua kotak ada: satu dengan kelereng biru dan yang lain dengan kelereng merah. Tanpa melihat kelereng yang kita ambil dari kotak, model kami diminta untuk menebak warnanya. Jika model salah menebak, misalnya bilang kita ambil kelereng merah padahal kita ambil kelereng biru, kita mencatat poin salah. Jika model menebak benar, misalnya bilang kita ambil kelereng biru, kita mencatat poin benar[16].

Metode untuk mencatat poin yang benar dan salah dikenal sebagai matriks kekacauan. Kita telah menulis semua kemungkinan jawaban yang benar dan yang salah di matriks ini. Dengan melihat catatan ini, kita dapat mengetahui seberapa baik model kita dalam menebak kelereng dan apa yang perlu diperbaiki untuk menjadikannya lebih pintar. Ada empat komponen penting dalam *matrix confusion*. Tabel berikut menggabungkan semua bagian tersebut. Tujuan utama dari *confusion matrix* adalah untuk memvisualisasikan dan menganalisis hasil prediksi yang dibuat oleh model. Ini memungkinkan kita untuk lebih mudah memahami kelebihan dan kekurangan model dalam mengklasifikasikan data[16].

Tabel 2.2 *Confusion Matrix*[16].

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Untuk memahami tabel di atas, kita akan menggunakan contoh kasus "COVID-19", di mana kita asumsikan kita membuat model klasifikasi menggunakan dataset COVID-19 yang diberi label atau class "positif" dan "negatif". Penjelasan keempat bagian dari confusion matrix yang ditampilkan di atas adalah sebagai berikut[16]:

1. True Positive (TP): merupakan jumlah tebakan yang benar, di mana model kita menebak seseorang terinfeksi COVID-19, dan kenyataannya memang benar. Misalnya, model menebak seseorang positif COVID-19, dan hasil tes memang menunjukkan positif COVID-19[16].
2. True Negative (TN): merupakan jumlah tebakan yang benar, di mana model kita menebak seseorang tidak terinfeksi COVID-19, dan kenyataannya memang benar. Misalnya, model menebak seseorang negatif COVID-19, dan hasil tes memang menunjukkan negatif COVID-19[16].
3. False Positive (FP): merupakan jumlah tebakan di mana model kita menebak seseorang terinfeksi COVID-19, tapi kenyataannya itu salah. Misalnya, model menebak seseorang positif COVID-19, padahal hasil tes sebenarnya menunjukkan negatif COVID-19[16].
4. False Negative (FN): merupakan jumlah tebakan di mana model kita menebak seseorang tidak terinfeksi COVID-19, tapi kenyataannya itu salah.

Misalnya, model menebak seseorang negatif COVID-19, padahal hasil tes sebenarnya menunjukkan positif COVID-19[16].

2.2.10 Penerapan Metode

Artificial Neural Network (ANN) telah banyak digunakan dalam kasus klasifikasi medis, termasuk klasifikasi penyakit tiroid. Misalnya, penelitian yang dilakukan oleh Jones et al. (2022) menggunakan ANN untuk mengklasifikasikan kondisi hipotiroid dan hipertiroid berdasarkan data pasien. Model ini sangat akurat dalam membedakan berbagai kelas tiroid, yang merupakan langkah penting dalam membantu diagnosis medis.

Contoh kasus : klasifikasi penyakit tiroid

Pada contoh ini, kita menggunakan dataset sederhana yang berisi beberapa variabel kesehatan (seperti kadar T3, T4, dan TSH) sebagai input, dan tujuannya adalah untuk mengklasifikasikan apakah pasien mengalami hipotiroid, hipertiroid, atau dalam kondisi normal. Misalnya, berikut adalah tiga variabel input dari tiga pasien:

PASIEN	T3	T4	TSH	KELAS
1	2.3	1.9	0.5	Normal
2	1.5	2.4	6.0	Hipotiroid
3	3.1	3.0	0.3	Hipertiroid

Tahap-tahap perhitungan manual ANN :

Inisialisasi Bobot dan Bias Misalkan kita memiliki tiga neuron di layer input dan satu neuron di layer output. Untuk setiap input, bobot awal (w_i) ditentukan secara acak. Sebagai contoh, bobot awal untuk input (T3, T4, TSH) adalah:

$$W = [0.1, 0.2, -0.1]$$

Bias awal ditentukan sebagai:

$$b = 0.05$$

Forward Propagation Forward propagation adalah tahap menghitung output dari model ANN berdasarkan input dan bobot yang diberikan. Pada neuron di lapisan output, nilai keluaran dihitung menggunakan fungsi aktivasi Sigmoid:

$$y = f(z) = \frac{1}{1 + e^{-z}}$$

Dimana $z = \sum(w_i \cdot x_i) + b$.

Misalkan untuk pasien 1:

$$z = (0.1 \cdot 2.3) + (0.2 \cdot 1.9) + (-0.1 \cdot 0.5) + 0.05$$

$$z = 0.23 + 0.38 - 0.05 + 0.05 = 0.61$$

$$y = \frac{1}{1 + e^{-0.61}} = 0.647$$

Evaluasi Loss (Fungsi Biaya) Misalkan target untuk pasien ini adalah "Normal" (label 0), dan kita menggunakan fungsi cross-entropy loss:

$$L = -[y \log(y) + (1 - y) \log(1 - y)]$$

$$L = -[0 \cdot \log(0.647) + (1 - 0) \cdot \log(1 - 0.647)]$$

$$L = -\log(0.353) = 1.041$$

Backpropagation dan pembaruan bobot, tahap backpropagation dilakukan untuk memperbaharui bobot dengan menggunakan gradien dari fungsi biaya terhadap setiap bobot.

Formula gradien untuk bobot adalah :

$$\Delta w_i = n \cdot \frac{\partial L}{\partial w_i}$$

Dimana n adalah laju pembelajaran.

Jika laju pembelajaran $n = 0.01$, dan kita hitung perubahan bobot (Δw) untuk bobot pertama:

$$\frac{\partial L}{\partial w_1} = (y - \hat{y}) \cdot x_1$$

Misalkan $y = 0$ dan $\hat{y} = 0.647$:

$$\frac{\partial L}{\partial w_1} = (0 - 0.647) \cdot 2.3 = -1.488$$

$$\Delta w_1 = 0.01 \cdot (-1.488) = -0.01488$$

Bobot baru w_1 menjadi :

$$w_1 = 0.1 - 0.01488 = 0.08512$$

Proses ini diulangi untuk setiap bobot dan bias.

Evaluasi Kinerja Model Setelah beberapa iterasi, model dilatih menggunakan data training, dan kinerjanya dievaluasi dengan data uji. Beberapa metrik evaluasi yang digunakan antara lain akurasi, sensitivitas, spesifisitas, dan AUC.

Contoh perhitungan akurasi manual sederhana: misalkan hasil prediksi untuk 3 pasien adalah sebagai berikut :

Pasien 1 (prediksi: Normal, Aktual:Normal) = Benar

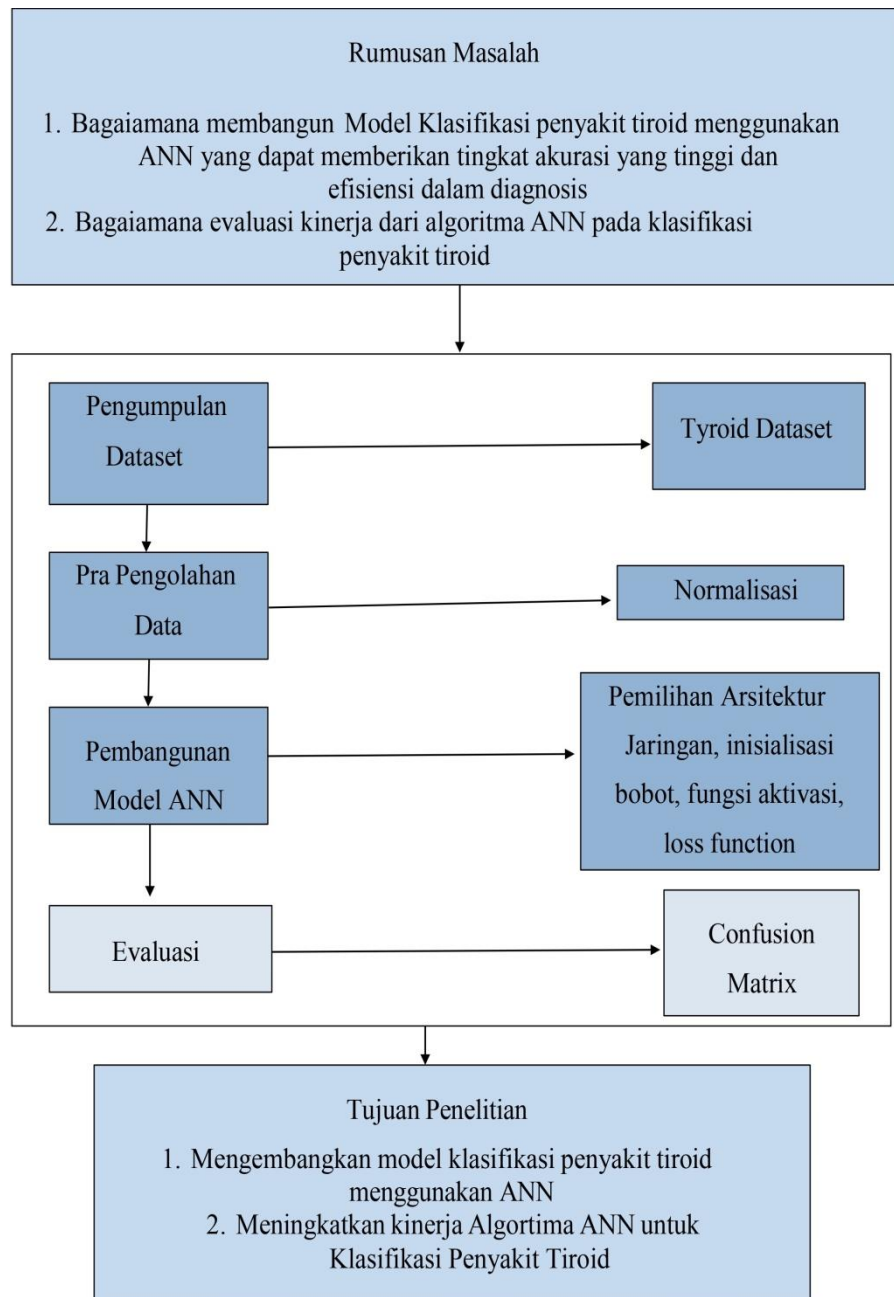
Pasien 2 (Prediksi: Hipotiroid, Aktual:Hipotiroid) = Benar

Pasien 3 (Prediksi: Hipertiroid, Aktual:Hipertiroid) = Benar

Akurasi:

$$\text{Akurasi} = \frac{\text{jumlah prediksi benar}}{\text{total data}} = \frac{3}{3} = 100\%$$

2.3 Kerangka Pikir



Gambar 2.8 Kerangka Pikir

BAB III

METODE PENELITIAN

3.1 Jenis, Metode, Subyek, Objek, Waktu, dan Lokasi Penelitian

Penelitian ini merupakan penelitian eksperimental ini menggunakan algoritma *Artificial Neural Network* (ANN) untuk membangun model klasifikasi penyakit tiroid. Penelitian ini termasuk dalam kategori penelitian kuantitatif karena melibatkan pengukuran akurasi model dan evaluasi kinerjanya menggunakan metrik yang diukur.

Penelitian ini menggunakan pembelajaran mesin (*machine learning*) dengan supervised learning. Data yang dikumpulkan akan digunakan untuk melatih dan menguji model ANN.

Model ANN yang akan dikembangkan untuk klasifikasi penyakit tiroid adalah subjek penelitian ini. Data pasien yang digunakan dalam penelitian ini berisi medis tentang penyakit tiroid, seperti kadar hormon tiroid (T3(*triiodothyronine*), T4(*thyroxine*), dan TSH(*thyroid stimulating hormone*)), serta variabel seperti usia, jenis kelamin, dan gejala klinis.

Penelitian direncanakan berlangsung dari September 2024 hingga Desember 2024. Mulai dari pengumpulan data, preprocessing, model pembangunan, implementasi, dan evaluasi kinerja. Penelitian ini dilakukan di Universitas Ichsan Gorontalo.

3.2 Pengumpulan Data

Data dalam penelitian ini merupakan data sekunder, dataset ini berisi informasi medis tentang pasien yang kadar hormon tiroidnya diperiksa (T3(*triiodothyronine*), T4(*thyroxine*), dan TSH(*thyroid stimulating hormone*)), yang akan digunakan untuk melatih dan menguji model ANN. Setiap baris dataset mewakili satu pasien, dan setiap fitur menggambarkan karakteristik klinis yang digunakan untuk memprediksi apakah pasien tersebut memiliki penyakit tiroid.

Tabel 3.1 Variabel Data

No	Nama Atribut	Type	Value	Ket
1	Usia	Numerik	Angka	Variabel Input
2	Jenis Kelamin	Kategori	Laki-Laki, Perempuan	Variabel Input
3	Riwayat Merokok	Kategori	Ya, Tidak	Variabel Input
4	Riwayat Radioterapi	Kategori	Ya, Tidak	Variabel Input
5	Fungsi Tiroid	Kategori	<i>Euthyroid</i>	Variabel Input
6	Pemeriksaan Fisik	Kategori	<i>Single Nodular(Goiter- Left, Goiter- Right, Multinodular Goiter</i>	Variabel Input
7	Adenopati	Kategori	Ya, Tidak	Variabel Input
8	Patologi	Kategori	Micropapillary	Variabel Input
9	Fokalitas	Kategori	<i>Uni-Focal, Multi-Focal</i>	Variabel Input
10	Risiko	Kategori	<i>Low</i>	Variabel Input
11	T(Tumor)	Kategori	T1a	Variabel Input
12	N(Kelenjar Getah Bening)	Kategori	NO	Variabel Input
13	M(Metasis)	Kategori	MO	Variabel Input
14	Stadium	Kategori	I	Variabel Input
15	Respons	Kategori	<i>Excellent, Inderminate</i>	Variabel Input
16	Kambuh	Kategori	Ya, Tidak	Variabel Output

3.3 Analisis Data

3.3.1 Pra Pengolahan

Dalam melakukan sejumlah tindakan penting untuk membersihkan dan mempersiapkan data sebelum digunakan oleh model *Artificial Neural Network* (ANN). Pra-pengolahan data ini sangat penting karena model ANN membutuhkan data yang seragam dan berkualitas tinggi untuk memberikan hasil terbaik. Pra-pengolahan data terdiri dari langkah-langkah berikut[17]:

1. Pembersihan Data (*Data Cleaning*) : Pastikan bahwa data yang digunakan bersih dan konsisten adalah langkah pertama dalam pra-pengolahan data. Data yang hilang atau tidak konsisten dapat menyebabkan masalah dalam model pelatihan[17].
2. Normalisasi : Setelah data dibersihkan, langkah berikutnya adalah normalisasi. Normalisasi adalah proses mengubah data ke dalam rentang tertentu, biasanya antara 0 dan 1, agar ANN dapat memproses data dengan lebih baik dan menghasilkan model yang lebih akurat. Hal ini dilakukan agar ANN dapat memproses data dengan lebih baik dan menghasilkan model yang lebih akurat[17].
3. *Feature Selection* (Pemilihan Fitur) : Dalam memilih fitur yang paling relevan untuk model ANN. Tidak semua fitur dalam dataset mungkin relevan atau berkontribusi signifikan terhadap prediksi ; Namun, memilih fitur yang tepat dapat meningkatkan performa model dan mengurangi waktu pelatihan. Proses ini disebut sebagai pengaturan fitur[17].

3.4 Pelatihan

3.4.2 Validasi data

Validasi dilakukan untuk memastikan model tidak terlalu sesuai dengan data pelatihan (*overfitting*). Teknik yang digunakan adalah:

1. *Cross-validation*: Dataset dibagi menjadi beberapa subset, dan model dilatih dan diuji pada subset yang berbeda untuk memastikan hasil yang konsisten[18].

2. Split Data: Dataset dibagi menjadi data pelatihan (biasanya 80%) dan data pengujian (20%). Data pengujian tidak digunakan selama pelatihan untuk mengukur kinerja model pada data yang tidak terlihat[18].

3.4.2 Evaluasi Model

Evaluasi model dilakukan untuk memahami seberapa baik model bekerja dalam hal prediksi yang akurat dan tepat. Beberapa metrik kinerja yang digunakan adalah:

1. Akurasi : Akurasi mengukur presentase prediksi yang benar dari semua prediksi yang dilakukan oleh model. Akurasi dapat dihitung dengan rumus[16] :

$$Akurasi = \frac{True\ Positives + True\ Negatives}{Total\ Samples}$$

2. *Sensitivitas* (Recall) : *Recall* adalah metrik yang mengukur kemampuan model untuk mendeteksi kasus positif dengan benar(misalnya, pasien dengan penyakit tiroid). Rumus untuk menghitung Recall[16] :

$$Recall = \frac{True\ Positives}{True\ positives + False\ Negatives}$$

3. *Spesifisitas* : Spesifisitas adalah metrik yang mengukur kemampuan model untuk mendeteksi kasus negatif dengan benar(misalnyam, pasien yang sehat). Rumus untuk menghitung Spesifisitas[16] :

$$Spesifisitas = \frac{True\ Negatives}{True\ Negatives + False\ Negatives}$$

4. *Area Under The Curve* (AUC) : AUC adalah metrik yang sering digunakan untuk masalah klasifikasi biner. AUC menggambarkan seberapa baik model dalam memisahkan antara kelas positif dan negatif. AUC dihasilkan dari ROC curve (Receiver Operating Characteristic Curve), yang menunjukkan hubungan antara True Positive Rate (Sensitivitas) dan False Positive Rate.

Skor AUC berkisar antara 0 hingga 1, dengan 1 menunjukkan performa sempurna (model dapat membedakan antara kelas positif dan negatif tanpa kesalahan), dan 0.5 menunjukkan performa acak (seperti melempar koin)[16].

BAB IV

HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Pengumpulan dataset pasien didapatkan dari dataset online yaitu dari KAGGLE sebanyak 383 record dalam bentuk file excel yang ditunjukkan seperti pada tabel berikut.

Tabel 4.1 Dataset Awal Pasien

No	Age	Gender	Smoking	Hx Smoking	Hx Radiothreapy	Thyroid Function	Physical Examination
1	27	F	No	No	No	Euthyroid	Single nodular goiter-left
2	34	F	No	Yes	No	Euthyroid	Multinodular goiter
3	30	F	No	No	No	Euthyroid	Single nodular goiter-right
4	62	F	No	No	No	Euthyroid	Single nodular goiter-right
5	62	F	No	No	No	Euthyroid	Multinodular goiter
6	52	M	Yes	No	No	Euthyroid	Multinodular goiter
7	41	F	No	Yes	No	hical Hyperthyroidi	Single nodular goiter-right
8	46	F	No	No	No	Euthyroid	Single nodular goiter-right
9	51	F	No	No	No	Euthyroid	Single nodular goiter-right
10	40	F	No	No	No	Euthyroid	Single nodular goiter-right
11	75	F	No	No	No	Euthyroid	Single nodular goiter-right
12	59	F	No	No	No	Euthyroid	Single nodular goiter-left
13	49	F	No	No	No	Euthyroid	Multinodular goiter
...							
383	67	M	Yes	No	No	Euthyroid	Multinodular goiter

Adenopathy	Pathology	Focality	Risk	T	N	M	Stage	Response	Recurred
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Indeterminate	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Multi-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Multi-Focal	Low	T1a	NO	M0	I	Indeterminate	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
No	Micropapillary	Uni-Focal	Low	T1a	NO	M0	I	Excellent	No
Bilateral	Papillary	Multi-Focal	High	T4b	N1b	M0	IVA	Structural Incomplete	Yes

Dari hasil pengumpulan data yang sudah dilakukan mendapatkan sejumlah 383 *record* data. Dengan jumlah variabel X sebanyak 16 dan variabel Y sebanyak 1 variabel :

Variabel X ;

1. *Age* : Usia Pasien

2. *Gender* : Jenis kelamin pasien, (F untuk perempuan, M untuk laki-laki)
3. *Smoking* : Status merokok pasien, apakah pasien seorang perokok atau tidak.
4. *Hx Smoking* : Riwayat merokok pasien (Ya atau Tidak).
5. *Hx Radiotherapy* : Riwayat radioterapi pasien, apakah pasien pernah menjalani pengobatan radioterapi (Ya atau Tidak).
6. *Thyroid Function* : Status fungsi tiroid pasien, misalnya Euthyroid berarti fungsi tiroid normal.
7. *Physical Examination* : Hasil pemeriksaan fisik pasien, misalnya ada pembesaran kelenjar tiroid (goiter) atau nodul yang terdeteksi.
8. *Adenopathy* : Apakah ada pembesaran kelenjar getah bening yang terdeteksi (Ya atau Tidak).
9. *Pathology* : Jenis patologi yang ditemukan pada pasien, misalnya Micropapillary (jenis kanker tiroid).
10. *Focality* : Menunjukkan apakah kondisi tersebut bersifat uni-fokal (terjadi pada satu area) atau multi-fokal (terjadi pada beberapa area).
11. *Risk* : Tingkat risiko kondisi pasien, misalnya risiko rendah (Low).
12. *T* : Ukuran dan penyebaran tumor (misalnya T1a berarti tumor kecil dan terbatas).
13. *N* : Keterlibatan kelenjar getah bening (misalnya N0 berarti tidak ada keterlibatan kelenjar getah bening).
14. *M* : Penyebaran metastasis (misalnya M0 berarti tidak ada keterlibatan kelenjar getah bening).
15. *Stage* : Tahapan Kanker secara keseluruhan, misalnya Stadium I, II, III, dan seterusnya.
16. *Response* : Respon pasien terhadap pengobatan, misalnya sangat baik (Excellent) atau tidak pasti (Indeterminate).

Variabel Y ;

1. *Recurred* : Menunjukkan apakah kondisi tersebut telah kambuh atau tidak (Ya atau Tidak).

4.2 Hasil Pengolahan Data

4.2.1 Pembagian data *training* dan data *testing*

Dataset dibagi menjadi dua bagian dengan perbandingan 80:20, yaitu 80% data untuk pelatihan (training) dan 20% data untuk pengujian (testing). Pembagian dilakukan dengan rumus[11]:

$$N_{train} = 0.8 \times N_{total}$$

$$N_{test} = 0.2 \times N_{total}$$

Dengan adalah jumlah total data dalam dataset. Berdasarkan rumus tersebut, hasil pembagian adalah:

- Total data : 383 sampel
- Data latih (training test) : $0.8 \times 383 = 306$ sampel
- Data uji (testing test) : $0.2 \times 383 = 77$ sampel

4.2.2 Normalisasi

Sangat penting untuk pembelajaran mesin, terutama untuk model berbasis gradien seperti Artificial Neural Network (ANN) dan Logistic Regression, normalisasi adalah proses penskalaan data agar berada dalam rentang tertentu, biasanya antara 0 dan 1, untuk memastikan bahwa semua fitur memiliki bobot yang seimbang dalam pemodelan[17].

Normalisasi dilakukan untuk :

- Menghindari bias fitur dengan skala besar
- Mempercepat kovergensi model
- Meningkatkan akurasi model

Rumus Normalisasi

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Di mana :

- X = Nilai Asli
- X min = Nilai minimum
- X max = nilai maksimum
- X scaled = Nilai setelah normalisasi

Tabel 4.2 Sampel data

Age	Gender	Smoking	Hx_Smoking	Hx_Radiotherapy	Thyroid_Function	Physical_Examination	Adenopathy
27	0	0	0	0	2	3	3
34	0	0	1	0	2	1	3
30	0	0	0	0	2	4	3
62	0	0	0	0	2	4	3
62	0	0	0	0	2	1	3

Pathology	Focality	Risk	T	N	M	Stage	Response
2	1	2	0	0	0	0	2
2	1	2	0	0	0	0	1
2	1	2	0	0	0	0	1
2	1	2	0	0	0	0	1
2	0	2	0	0	0	0	1

Perhitungan Normalisasi :

Sebagai contoh, jika fitur “Age” memiliki rentang nilai asli antara 20-80, dan X =50, maka normalisasinya dihitung sebagai berikut :

$$X_{scaled} = \frac{50 - 20}{80 - 20} = \frac{30}{60} = 0.5$$

Hasil dari normalisasi beberapa sampel ditunjukkan dalam tabel berikut :

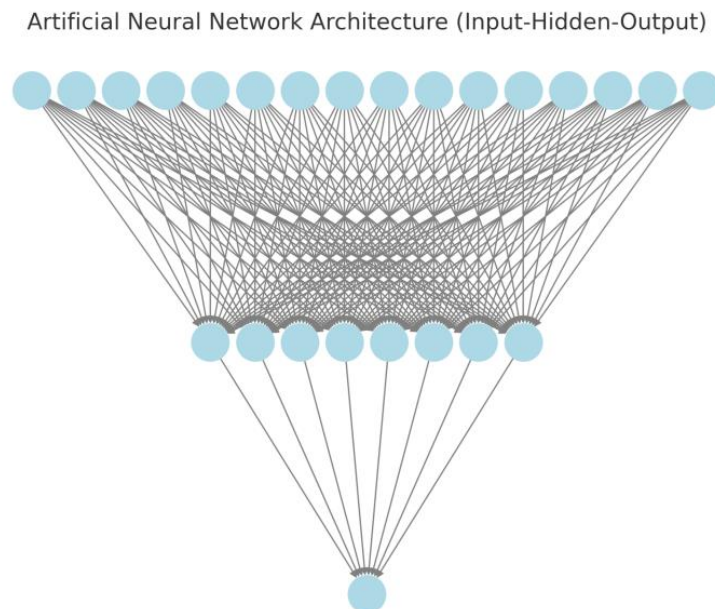
Tabel 4.3 Hasil Normalisasi

Age	Gender	Smoking	Hx_Smoking	Hx_Radiotherapy	Thyroid_Function	Physical_Examination	Adenopathy
0.537	0	0	0	0	0.5	0.75	0.6
0.641	0	0	0	0	1	0.25	0.2
0.194	0	0	0	0	0.5	0.75	0.6
0.298	0	0	0	0	0.5	0.25	0.2
0.731	0	0	1	0	0.5	0.25	0.6

Pathology	Focality	Risk	T	N	M	Stage	Response
0	1	0.5	0.5	0	0	0	1
1	0	0.5	0.67	1	0	0.25	1
1	1	1	0.33	0	0	0	0.333
1	0	0	0.83	1	0	0	1
0	0	0	0.83	0	1	1	1

4.2.3 Proses dengan ANN

Artificial Neural Network (ANN) menggunakan otak manusia sebagai model pembelajaran mesin. Untuk meningkatkan akurasi model, setiap neuron dalam ANN menerima bobot dan bias yang diperbarui selama proses pelatihan. ANN terdiri dari lapisan neuron yang saling terhubung yang berfungsi untuk melakukan pemetaan antara input dan output. Berikut proses arsitektur dari Artificial Neural Network berdasarkan dari hasil pengumpulan data yang sudah dilakukan sebelumnya menghasilkan 16 variabel inputan (X) dengan jumlah hidden layer 8 serta output (Y) [5]:



Gambar 4.1 Arsitektur ANN[5].

1. Inisialisasi Bobot dan Bias

Dalam jaringan syaraf tiruan (Artificial Neural Network/ANN), bobot (weights) dan bias adalah dua komponen utama yang menentukan bagaimana input diproses untuk menghasilkan output.

Bobot dan bias merupakan parameter penting dalam jaringan neural. Bobot (weights) digunakan untuk menentukan seberapa besar pengaruh suatu input terhadap output, sedangkan bias adalah nilai tambahan yang memungkinkan model untuk melakukan pergeseran terhadap fungsi aktivasi, sehingga meningkatkan fleksibilitas dan akurasi prediksi. Pada tahap inisialisasi, nilai bobot dan bias biasanya dipilih secara acak (random) dalam rentang nilai kecil, umumnya kurang dari 1. Pemilihan nilai awal secara acak ini bertujuan agar proses pelatihan (training) model dapat dimulai tanpa memberikan preferensi awal tertentu, sekaligus menghindari simetri dalam pembaruan parameter saat proses backpropagation berlangsung.

Tabel 4.4 Inisialisasi Bobot dan Bias

Feature	W1	W2	W3	W4	W5	W6	W7	W8
Age	0.3	0.9	0.7	0.5	0.1	0.1	0.0	0.8
Gender	0.6	0.7	0.0	0.9	0.8	0.2	0.1	0.1
Smoking	0.3	0.5	0.4	0.2	0.6	0.1	0.2	0.3
Hx_Smoking	0.4	0.7	0.1	0.5	0.5	0.0	0.6	0.1
Hx_Radiotherapy	0.0	0.9	0.9	0.8	0.3	0.0	0.6	0.4
Thyroid_Function	0.1	0.4	0.0	0.9	0.2	0.6	0.3	0.5
Physical_Examination	0.5	0.1	0.9	0.7	0.9	0.8	0.5	0.9
Adenopathy	0.0	0.1	0.0	0.3	0.3	0.2	0.8	0.3
Pathology	0.2	0.5	0.1	0.8	0.0	0.9	0.7	0.1
Focality	0.0	0.8	0.7	0.7	0.7	0.0	0.3	0.1
Risk	0.8	0.6	0.3	0.0	0.3	0.3	0.7	0.6
T	0.8	0.4	0.1	0.7	0.7	0.5	0.7	0.4
N	0.5	0.4	0.0	0.1	0.0	0.6	0.3	0.5
M	0.9	0.2	0.4	0.7	0.2	0.0	0.2	0.1
Stage	0.9	0.8	0.6	0.8	0.8	0.1	0.8	0.5
Response	0.8	0.8	0.3	0.1	0.2	0.4	0.8	0.8

Bias hidden layer :

N-1 : 0.0

N-2 : 0.5

N-3 : 0.4

N-4 : 0.2

N-5 : 0.1

N-6 : 0.3

N-7 : 0.9

N-8 : 0.3

Bobot *hidden* ke *output* layer :

N-1 : 0.5

N-2 : 0.7

N-3 : 0.3

N-4 : 0.9

N-5 : 0.9

N-6 : 0.2

N-7 : 0.4

N-8 : 0.3

Bias *output* layer : N-1 : 0.2

2. Forward Propagation

Menghitung output hidden layer menggunakan fungsi sigmoid sebagai berikut :

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- Contoh Input vector baris pertama (x1)

0,537	0	0	0	0	0,5	0,75	0,6
0	1	0,5	0,5	0	0	0	1

Neuron 1 :

$$Z_1 = (W_1 \cdot x) + b_1$$

$$\begin{aligned} Z_1 = & (0,3 \times 0,537) + (0,6 \times 0) + (0,3 \times 0) + (0,4 \times 0) + (0,0 \times 0) + (0,1 \times 0,5) \\ & + (0,5 \times 0,75) + (0,0 \times 0,6) + (0,2 \times 0) + (0,0 \times 1) + (0,8 \times 0,5) \\ & + (0,8 \times 0,5) + (0,5 \times 0) + (0,9 \times 0) + (0,9 \times 0) + (0,8 \times 1) \end{aligned}$$

$$\begin{aligned} Z_1 = & 0,1611 + 0 + 0 + 0 + 0 + 0,05 + 0,375 + 0 + 0 + 0 + 0,4 + 0 + 0 + 0 \\ & + 0,8 + 0,0 = 2,1861 \end{aligned}$$

$$h_1 = \sigma(Z_1) = \frac{1}{1 + e^{-2,1861}} = 0,898$$

Neuron 2 :

$$Z_2 = (W_2 \cdot x) + b_2$$

$$\begin{aligned} Z_2 = & (0,9 \times 0,537) + (0,7 \times 0) + (0,5 \times 0) + (0,7 \times 0) + (0,9 \times 0) \\ & + (0,4 \times 0,5) + (0,1 \times 0,75) + (0,1 \times 0,6) + (0,5 \times 0) + (0,8 \times 1) \\ & + (0,6 \times 0,5) + (0,4 \times 0,5) + (0,4 \times 0) + (0,2 \times 0) + (0,8 \times 0) \\ & + (0,8 \times 1) \end{aligned}$$

$$\begin{aligned} Z_2 = & 0,4833 + 0 + 0 + 0 + 0 + 0,2 + 0,75 + 0,06 + 0 + 0,8 + 1,1 + 0,9 + 0 \\ & + 0 + 0 + 0,8 + 0,5 = 4,9183 \end{aligned}$$

$$h_2 = \sigma(Z_2) = \frac{1}{1 + e^{-4,9183}} = 0,993$$

Neuron 3 :

$$Z_3 = (W_3 \cdot x) + b_3$$

$$\begin{aligned} Z_3 = & (0,7 \times 0,537) + (0,0 \times 0) + (0,4 \times 0) + (0,1 \times 0) + (0,9 \times 0) + (0,0 \times 0,5) \\ & + (0,9 \times 0,75) + (0,0 \times 0,6) + (0,1 \times 0) + (0,7 \times 1) + (0,3 \times 0,5) \\ & + (0,1 \times 0,5) + (0,0 \times 0) + (0,4 \times 0) + (0,6 \times 0) + (0,3 \times 1) \\ Z_3 = & 0,3759 + 0 + 0 + 0 + 0 + 0 + 0 + 1,65 + 0 + 0 + 0,7 + 0,15 + 0,5 + 0 + 0 + 0 \\ & + 0,3 + 0,4 = 3,6259 \end{aligned}$$

$$h_3 = \sigma(Z_3) = \frac{1}{1 + e^{-3,6259}} = 0,974$$

Neuron 4 :

$$Z_4 = (W_4 \cdot x) + b_4$$

$$\begin{aligned} Z_4 = & (0,5 \times 0,537) + (0,9 \times 0) + (0,2 \times 0) + (0,5 \times 0) + (0,8 \times 0) + (0,9 \times 0,5) \\ & + (0,7 \times 0,75) + (0,3 \times 0,6) + (0,8 \times 0) + (0,7 \times 1) + (0,0 \times 0,5) \\ & + (0,7 \times 0,5) + (0,1 \times 0) + (0,7 \times 0) + (0,8 \times 0) + (0,1 \times 1) \\ Z_4 = & 0,2685 + 0,45 + 0,525 + 0,18 + 0,7 + 0,35 + 0,1 + 0,2 = 2,7735 \end{aligned}$$

$$h_4 = \sigma(Z_4) = \frac{1}{1 + e^{-2,7735}} = 0,941$$

Neuron 5 :

$$Z_5 = (W_5 \cdot x) + b_5$$

$$\begin{aligned} Z_5 &= (0,1 \times 0,537) + (0,8 \times 0) + (0,6 \times 0) + (0,5 \times 0) + (0,3 \times 0) + (0,2 \times 0,5) \\ &\quad + (0,9 \times 0,75) + (0,3 \times 0,6) + (0,0 \times 0) + (0,7 \times 1) + (0,3 \times 0,5) \\ &\quad + (0,7 \times 0,5) + (0,0 \times 0) + (0,2 \times 0) + (0,8 \times 0) + (0,2 \times 1) \\ Z_5 &= 0,0537 + 0,1 + 0,675 + 0,18 + 0,7 + 0,15 + 0,35 + 0,2 + 0,1 = 2,5087 \end{aligned}$$

$$h_5 = \sigma(Z_5) = \frac{1}{1 + e^{-2,5087}} = 0,923$$

Neuron 6 :

$$Z_6 = (W_6 \cdot x) + b_6$$

$$\begin{aligned} Z_6 &= (0,1 \times 0,537) + (0,2 \times 0) + (0,1 \times 0) + (0,0 \times 0) + (0,0 \times 0) + (0,6 \times 0,5) \\ &\quad + (0,8 \times 0,75) + (0,2 \times 0,6) + (0,9 \times 0) + (0,0 \times 1) + (0,3 \times 0,5) \\ &\quad + (0,5 \times 0,5) + (0,6 \times 0) + (0,0 \times 0) + (0,1 \times 0) + (0,4 \times 1) \\ Z_6 &= 0,0537 + 0,3 + 0,6 + 0,12 + 0,15 + 0,25 + 0,4 + 0,3 = 2,1737 \end{aligned}$$

$$h_6 = \sigma(Z_6) = \frac{1}{1 + e^{-2,1737}} = 0,896$$

Neuron 7 :

$$Z_7 = (W_7 \cdot x) + b_7$$

$$\begin{aligned} Z_7 &= (0,0 \times 0,537) + (0,1 \times 0) + (0,2 \times 0) + (0,6 \times 0) + (0,6 \times 0) + (0,3 \times 0,5) \\ &\quad + (0,5 \times 0,75) + (0,8 \times 0,6) + (0,7 \times 0) + (0,3 \times 1) + (0,7 \times 0,5) \\ &\quad + (0,7 \times 0,5) + (0,3 \times 0) + (0,2 \times 0) + (0,8 \times 0) + (0,8 \times 1) \\ Z_7 &= 0,15 + 0,375 + 0,48 + 0,3 + 0,35 + 0,35 + 0,8 + 0,4 = 3,205 \end{aligned}$$

$$h_7 = \sigma(Z_7) = \frac{1}{1 + e^{-3,205}} = 0,9600$$

Neuron 8 :

$$Z_8 = (W_8 \cdot x) + b_8$$

$$\begin{aligned}
Z_7 &= (0,8 \times 0,537) + (0,1 \times 0) + (0,3 \times 0) + (0,1 \times 0) + (0,4 \times 0) + (0,5 \times 0,5) \\
&\quad + (0,9 \times 0,75) + (0,3 \times 0,6) + (0,1 \times 0) + (0,1 \times 1) + (0,6 \times 0,5) \\
&\quad + (0,4 \times 0,5) + (0,5 \times 0) + (0,1 \times 0) + (0,5 \times 0) + (0,8 \times 1) \\
Z_7 &= 0,4296 + 0,25 + 0,675 + 0,18 + 0,1 + 0,3 + 0,2 + 0,8 + 0,3 = 3,2346 \\
h_7 &= \sigma(Z_7) = \frac{1}{1 + e^{-3,2346}} = 0,9627
\end{aligned}$$

Menghitung output layer :

$$Z_o = W_{01} \cdot h_1 + W_{02} \cdot h_2 + W_{03} \cdot h_3 + W_{04} \cdot h_4 + W_{05} \cdot h_5 + W_{06} \cdot h_6 + W_{07} \cdot h_7 + W_{08} \cdot h_8 + b_o$$

$$\begin{aligned}
Z_o &= (0,5 \times 0,898) + (0,7 \times 0,993) + (0,3 \times 0,974) + (0,9 \times 0,941) \\
&\quad + (0,9 \times 0,923) + (0,2 \times 0,896) + (0,4 \times 0,9600) \\
&\quad + (0,3 \times 0,9627) + 0,2
\end{aligned}$$

$$\begin{aligned}
Z_o &= 0,449 + 0,6951 + 0,2922 + 0,8307 + 0,1792 + 0,384 + 0,28881 + 0,2 \\
&= 4,16591
\end{aligned}$$

$$O = \sigma(Z_o) = \frac{1}{1 + e^{-4,16591}} = 0,983$$

Output akhir = 0,983

Output ini menunjukan kemungkinan bahwa Recurred adalah 1, dengan ambang di atas >0,5, sehingga model dapat mengklasifikasikan 1 (positif) atau kambuh.

Proses ini dilakukan secara iteratif atau berulang-ulang pada data yang sudah ada, dikenal sebagai epoch. Pada setiap epoch, proses tersebut menghitung eror, mengubah bobot, dan bias untuk mendapatkan model yang paling efisien.

4.3 Confusion Matrix

Berdasarkan hasil yang diperoleh dari confusion matrix, berikut interpretasi nilai yang ada, seperti pada tabel berikut.

Tabel 4.5 Testing.

	Klasifikasi (Tidak) 0	Klasifikasi (Ya) 1
Nilai Aktual (Tidak) 0	58	0

Nilai Aktual (Ya) 1	2	17
---------------------	---	----

Berdasarkan *confusion matrix*, maka dapat dihitung *metrix* evaluasi secara manual :

Berikut keterangan dari hasil *confusion matrix* :

- *True Positive* (TP) = 58 (Klasifikasi “Ya” benar)
- *False Positive* (FP) = 0 (Klasifikasi “Ya” salah)
- *True Negative* (TN) : 17 (Klasifikasi “Tidak” benar)
- *False Negative* (FN) : 2 (Klasifikasi “Tidak” salah)

1). *Accuracy*

Accuracy berfungsi untuk mengukur proporsi klasifikasi dengan benar (positif dan negatif) dibandingkan dengan total klasifikasi.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{58 + 17}{58 + 17 + 0 + 2} = \frac{75}{77} = 0,974$$

2). *Precision* (Presisi)

Precision berfungsi untuk mengukur akurasi klasifikasi positif[11].

$$Precision = \frac{TP}{TP + FP} = \frac{58}{58 + 0} = \frac{58}{58} = 1$$

3). *Recall* (Sensitivitas)

Recall berfungsi untuk mengukur kemampuan model menemukan setiap kasus positif sebelumnya[11].

$$Recall = \frac{TP}{TP + FN} = \frac{58}{58 + 2} = \frac{58}{60} = 0,966$$

4). *F1-Score*

F1-Score berfungsi untuk dapat mengetahui rata-rata harmonik dari *precision* dan *recall*[11].

$$F1 - Score = 2X \frac{Precision \times Recall}{Precision + Recall} = 2X \frac{100 \times 0,966}{100 + 0,966} = \frac{96,6}{100,966} = 0,950$$

Hasil Perhitungan :

Accuracy = 0,974 atau 97,45 dibulatkan menjadi 97%

Precision = 100%

Recall = 0,966 atau 96,6% dibulatkan menjadi 97%

F1-Score = 0,950 atau 95,0% dibulatkan menjadi 95 %

Classification Report Metrics menunjukkan hasil metrik klasifikasi, yaitu skor precision, recall, dan f1 untuk dua kelas (0 dan 1), serta nilai agregat seperti ketepatan, nilai tengah makro, dan nilai tengah berat. Model ini menunjukkan performa yang baik dengan ketepatan sebesar 97%, yang menunjukkan ketepatan dalam mengklasifikasikan data. Skor precision dan recall untuk kedua kelas hampir seimbang, yang menunjukkan bahwa model tidak hanya tepat dalam klasifikasi positif tetapi juga dapat mengidentifikasi sebagian besar data positif sehingga membuat model ini efektif dalam hal klasifikasi.

BAB V

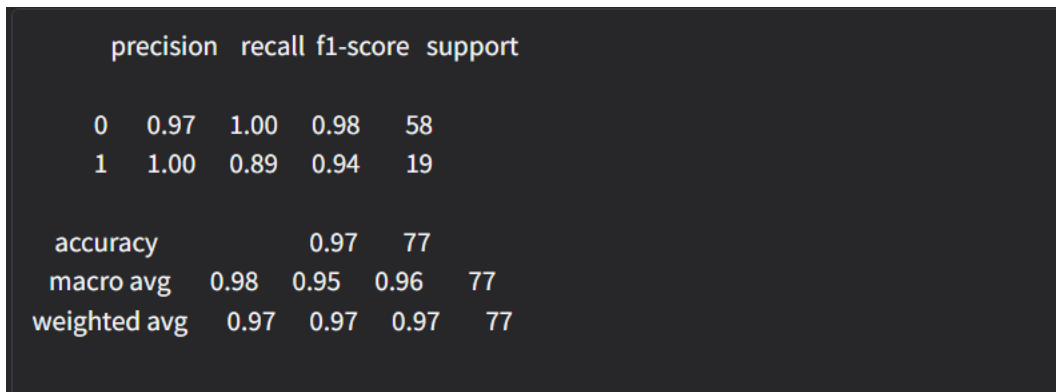
PEMBAHASAN

5.1 Pembahasan Pengumpulan Data

Pengumpulan data dilakukan secara online atau data publik yang diperoleh dari Kaggle. Data yang berhasil diperoleh berjumlah 383 entri, Kumpulan data ini yang dijadikan objek dalam mengembangkan pemodelan dalam pengujian kinerja algoritma *ANN*.

5.2 Pembahasan Hasil Pemodelan

Setelah dilakukan pemodelan dengan mengklasifikasikan penyakit tiroid menghasilkan akurasi yg sangat baik, yaitu dengan nilai **97%**. Dalam penelitian ini, kinerja model klasifikasi dinilai menggunakan metrik *precision*, *recall*, dan *f1-score*. Hasilnya menunjukkan bahwa model cukup akurat dalam memprediksi kedua kelas tersebut, dengan *precision* kelas 0 sebesar 0.97 dan *precision* kelas 1 sebesar 1.00. Untuk *recall*, model mengenali semua data kelas 0 dengan *recall* 1.00, tetapi *recall* kelas 1 sedikit lebih rendah, yaitu 0.89. Seperti yang ada pada gambar berikut.



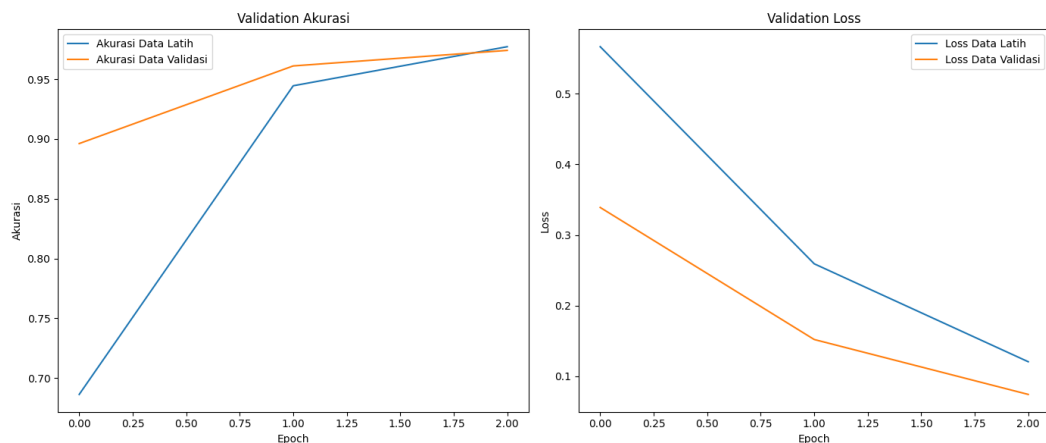
	precision	recall	f1-score	support
0	0.97	1.00	0.98	58
1	1.00	0.89	0.94	19
accuracy			0.97	77
macro avg	0.98	0.95	0.96	77
weighted avg	0.97	0.97	0.97	77

Gambar 5.1 Hasil Pemodelan.

5.3 Validation Akurasi dan Validation Loss

Dalam grafik *validation akurasi* dan *validation loss* menunjukkan bagaimana model berkembang selama proses pelatihan. Akurasi validasi ditunjukkan pada grafik pertama. Garis biru menunjukkan peningkatan yang konsisten dalam akurasi data latih, dan garis oranye menunjukkan peningkatan

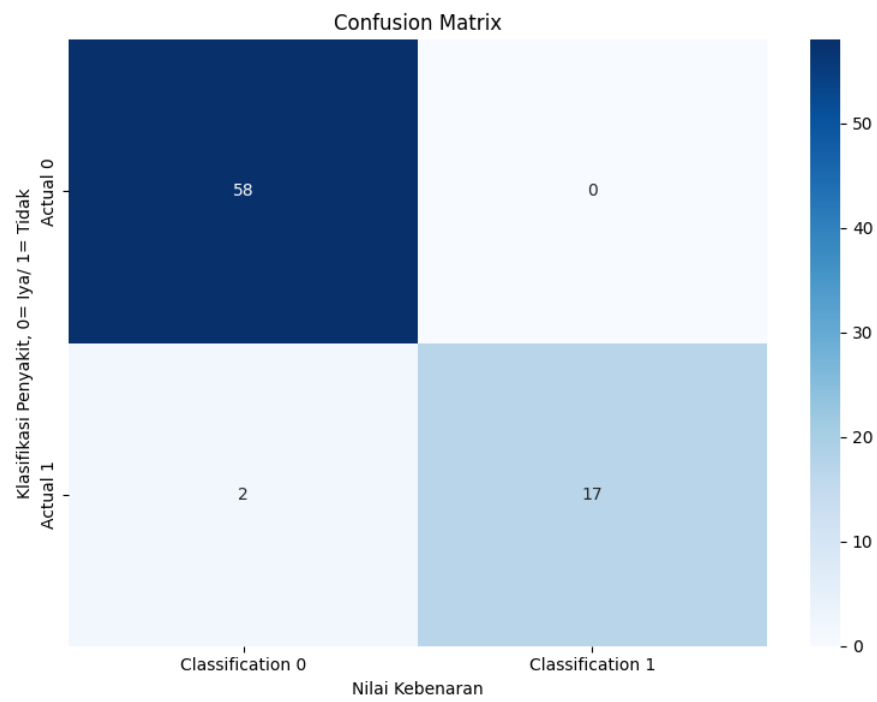
akurasi data validasi, meskipun dengan laju yang lebih lambat. Hal ini menunjukkan bahwa seiring berjalannya waktu, model menjadi semakin akurat dalam memprediksi data latih dan validasi. Dalam grafik kedua, kehilangan data latih (garis biru) terus menurun, menunjukkan bahwa model semakin baik dalam meminimalkan kesalahan pada data latih, dan kehilangan data validasi (garis oranye) juga menurun, menunjukkan bahwa model dapat menggeneralisasi dengan baik pada data yang tidak terlihat. Secara keseluruhan, grafik ini menunjukkan bahwa model memiliki kemampuan belajar yang baik dan mengoptimalkan akurasi serta mengurangi kesalahan pada data validasi. Seperti yang ada pada gambar berikut.



Gambar 5.2 Validation Akurasi dan Validation Loss.

5.4 Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi hasil klasifikasi model, Matriks ini menunjukkan jumlah data yang diklasifikasikan dengan benar maupun salah untuk masing-masing kelas. Ini menunjukkan bahwa model berhasil mengklasifikasikan data dengan benar. Pada baris pertama, yang mewakili kelas 0 (Tidak), terdapat 58 data yang benar diklasifikasikan sebagai kelas 0 (*True Negative*), dan 0 data yang salah diklasifikasikan sebagai kelas 1. Pada baris kedua, yang mewakili kelas 1 (Ya), terdapat 2 data yang sebenarnya kelas 1 tetapi salah diklasifikasikan sebagai kelas 0 (*False Negative*), dan 17 data yang benar diklasifikasikan sebagai kelas 1. Ini menunjukkan bahwa model berhasil mengklasifikasikan data dengan benar. Seperti yang ada pada gambar berikut.



Gambar 5.3 Confusion Matrix.

BAB VI

PENUTUP

6.1 Kesimpulan

Model *Artificial Neural Network (ANN)* yang dikembangkan dalam penelitian ini berhasil menunjukkan kinerja yang sangat baik dengan akurasi mencapai 97%, serta peningkatan akurasi dan penurunan loss yang signifikan seiring berjalannya *epoch*. Hasil evaluasi menunjukkan bahwa model ini dapat mengklasifikasikan kedua kelas (Ya dan Tidak) dengan memuaskan, seperti yang terlihat pada hasil *Confusion Matrix*. Namun, keterbatasan penelitian ini terletak pada jumlah dan keragaman dataset yang terbatas, yang dapat memengaruhi kemampuan *generalisasi* model.

6.2 Saran

Untuk penelitian selanjutnya, disarankan agar menggunakan dataset yang lebih besar dan lebih beragam agar model dapat diuji pada berbagai variasi data medis dan sumber yang berbeda. Hal ini bertujuan untuk menghindari *overfitting* dan meningkatkan akurasi serta kemampuan *generalisasi* model. Selain itu, perbandingan dengan algoritma lain seperti *Random Forest* atau *Support Vector Machine (SVM)* dapat dilakukan untuk mengetahui keunggulan masing-masing algoritma dalam klasifikasi p enyakit tiroid.

DAFTAR PUSTAKA

- [1] D. Pittara, “penyakit tiroid.” [Online]. Available: <https://www.alodokter.com/penyakit-tiroid>, diakses 14/09/2024.
- [2] D. KEMAS, E. Wianto, and M. TJANDRA, “Perancangan Buku Panduan Interaktif Mengenai Pengenalan Penyakit Gangguan Kelenjar Tiroid Pada Anak-Anak,” *Serat Rupa J. Des.*, vol. 2, no. 1, p. 40, 2018, doi: 10.28932/srjd.v2i1.476.
- [3] S. Y. Mustika and A. Wahyuni, “Manajemen Anestesi pada Papillary Thyroid Carcinoma: Sebuah Laporan Kasus,” *J. Penelit. Perawat Prof.*, vol. 4, no. 4, pp. 1085–1092, 2022.
- [4] C. Untuk and K. Harga, “Perbandingan Kinerja Algoritma Naive Bayes dan,” vol. 7, no. 1, pp. 20–24, 2020.
- [5] M. P. K. Tommy Dwi Putra, Ema Utami, “Klasifikasi penderita kanker Paru Paru Menggunakan Algoritma Artificial Neural Network (ANN),” *J. Explor. Univ. Teknol. Mataram*, 2022, [Online]. Available: <https://ojs.utmmataram.ac.id/index.php/explore/article/view/568>, diakses 10/10/2024.
- [6] D. G. Pradana, S. , Muhammad Luthfi Alghifari, Muhammad Farhan Juna, and D. Palaguna, “Klasifikasi Penyakit Jantung Menggunakan Metode Artificial Neural Network,” *Indones. J. Data Sci.*, 2022.
- [7] ander sriwi sri sucaty, murianto, “Penerapan Algoritma Artificial Neural Network Untuk Memprediksi Penyakit Gagal Jantung,” *Pola Kemitraan Pentahelix Dalam Pengemb. Desa Wisata Buwun Sejati, Lomb. Barat Ntb*, vol. 3, no. 4, pp. 413–446, 2024.
- [8] M. Resha, S. Syamsu, Andryanto, and Bakry, “Prediksi Penyebaran Kasus Tuberkulosis dengan Metode Artificial Neural Network dan Multi-Layer Perceptron di Kota Makassar,” *JNSTA ADPERTISI J.*, vol. 2, no. 1, pp. 16–21, 2022.
- [9] A. Alzahrani, A. S., & Alswailem, “Diagnosis of Thyroid Diseases Using AI-Based Systems: A Review,” *J. Healthc. Eng.*, 2021.

- [10] A. Alzahrani, A. S., & Alswailem, “Diagnosis of Thyroid Diseases Using AI-Based Systems,” *J. Healthc. Eng.*, 2021.
- [11] F. A. Hermawati, *Data Mining*. Yogyakarta: Andi, 2013.
- [12] M. Riadi, “Pengertian, Fungsi, Proses dan Tahapan Data Mining.” [Online]. Available: <https://www.kajianpustaka.com/2017/09/data-mining.html>, diakses 30/09/2024.
- [13] D. Armiady, “Analisis metode DBSCAN (Density-Based Spatial Clustering of Application with Noise) dalam mendeteksi data outlier,” *JURIKOM (Jurnal Ris. Komputer)*, 2022.
- [14] M. Yunus, “Jaringan Syaraf Tiruan (JST) — Bagian 1 (Pengenalan).” [Online]. Available: [file:///C:/Users/hp/Downloads/%236 Jaringan Syaraf Tiruan \(JST\) — Bagian 1 \(Pengenalan\) _ oleh Muhammad Yunus _ Medium.html](file:///C:/Users/hp/Downloads/%236%20Jaringan%20Syaraf%20Tiruan%20(JST)%20—%20Bagian%201%20(Pengenalan)%20_oleh%20Muhammad%20Yunus%20_Medium.html), diakses 08/10/2024.
- [15] Kantinit, “Belajar Jaringan Syaraf Tiruan (JST): Pengertian, Arsitektur, Cara Kerja dan Jenis-Jenisnya.” [Online]. Available: [Belajar Jaringan Syaraf Tiruan \(JST\)_ Arsitektur, Cara Kerja dan Jenis-Jenisnya.html](#), diakses 12/10/2024.
- [16] Rina, “Memahami Confusion Matrix: Accuracy, Precision, Recall, Specificity, dan F1-Score untuk Evaluasi Model Klasifikasi.” [Online]. Available: <https://esairina.medium.com/memahami-confusion-matrix-accuracy-precision-recall-specificity-dan-f1-score-610d4f0db7cf>, diakses 12/10/2024.
- [17] S. K. A. Firdausi, “4 Tahap Preprocessing Data, Beserta Penjelasan & Studi Kasus.” [Online]. Available: [https://dibimbing.id/blog/detail/mengenal-apa-itu-tahap-preprocessing-data#:~:text=Data preprocessing adalah proses persiapan dan transformasi data,berkualitas dan reliabel untuk keperluan analisis lebih lanjut](https://dibimbing.id/blog/detail/mengenal-apa-itu-tahap-preprocessing-data#:~:text=Data%20preprocessing%20adalah%20proses%20persiapan%20dan%20transformasi%20data,berkualitas%20dan%20reliabel%20untuk%20keperluan%20analisis%20lebih%20lanjut), diakses 12/10/2024.
- [18] K. S. Nugroho, “Validasi Model Klasifikasi Machine Learning pada RapidMiner.” [Online]. Available: <https://ksnugroho.medium.com/validasi-model-machine-learning-pada-rapidminer-50be0080df14>, diakses 12/10/2024

- [19] Dr. Suyanto, *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*. Bandung: Informatika, 2019.

DATA YANG DIKUMPULKAN

	Age	Gender	Smoking	Hx Smoking	Hx Radiotherapy	Thyroid Function	Physical Examination	Adeno pathy	Pathology	Focality	Risk	T	N	M	Stage	Response	Recurred
27	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
34	F	No	Yes	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
30	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
62	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
62	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Excellent	No
52	M	Yes	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Indeterminate	No
41	F	No	Yes	No	No	Clinical Hypothyroidism	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
46	F	No	No	No	Euthyroid		Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
51	F	No	No	No	Euthyroid		Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
40	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
75	F	No	No	No	No	Euthyroid	Single nodular goiter goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
59	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
49	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
50	F	No	No	No	No	Clinical Hypothyroidism	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
76	F	No	No	No	Euthyroid		Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
42	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
40	F	No	Yes	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
44	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Indeterminate	No
43	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
52	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Indeterminate	No
41	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
44	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
36	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
70	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
60	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Excellent	No
33	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
43	M	No	No	No	No	Subclinical Hypertyroidism	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
26	M	No	No	No	Euthyroid		Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
41	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
37	F	No	Yes	No	No	ubclinical Hypothyroidisr	Normal	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
37	F	No	No	No	Euthyroid		Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
30	F	No	No	No	No	Clinical Hypothyroidism	Normal	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
36	F	No	No	No	Euthyroid		Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
55	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
52	M	Yes	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
37	F	No	No	No	No	ubclinical Hypothyroidisr	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
31	F	No	No	No	No	Clinical Hypothyroidism	Diffuse goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
43	M	No	No	No	Euthyroid		Multinodular goiter	No	Micropapillary	Multi-Focal	Low	T1a	N0	M0	I	Excellent	No
40	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
34	F	No	No	No	No	Euthyroid	Diffuse goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
51	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
45	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
20	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
38	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
38	F	No	No	No	No	Clinical Hypothyroidism	Diffuse goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent	No
33	F	No	No	No	No		Euthyroid	Multinodular goiter	No	Micropapillary	Uni-Focal	Low	T1a	N0	M0	I	Excellent
31	F	No	No	No	No	Clinical Hypothyroidism	Multinodular goiter	No	Micropapillary	Multi-Focal	Intermediate	T1a	N1b	M0	I	Excellent	No
31	F	No	No	No	Euthyroid		Multinodular goiter	Normal	extensiv Papillary	Uni-Focal	Intermediate	T1a	N1b	M0	I	Actual Incomple	Yes
29	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Follicular	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
43	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Micropapillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
30	F	No	No	No	No	Subclinical Hypertyroidism	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Excellent	No
25	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Excellent	No
27	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
25	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
21	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Indeterminate	No
43	F	No	Yes	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
23	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
23	F	No	No	No	No	Clinical Hypothyroidism	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
43	F	No	No	No	Euthyroid		Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Indeterminate	No
24	M	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
35	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
54	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
54	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Indeterminate	No
22	F	No	No	No	No	Subclinical Hypertyroidism	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
38	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Excellent	No
35	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
51	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
22	M	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Excellent	No
51	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
40	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
69	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
31	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
29	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
30	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
28	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
22	F	No	No	No	No	Clinical Hypothyroidism	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
29	F	No	No	No	No		Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent
35	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
50	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Excellent	No
27	F	No	No	No	No	Clinical Hypothyroidism	Diffuse goiter	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
17	F	No	Yes	No	No		Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent
27	F	No	No	No	No	Subclinical Hypertyroidism	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
40	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Intermediate	T1b	N0	M0	I	Excellent	No
33	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Papillary	Uni-Focal	Intermediate	T1b	N1b	M0	I	Indeterminate	No
25	F	No	Yes	Yes	Yes	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
73	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Papillary	Uni-Focal	Low	T1b	N0	M0	I	Excellent	No
36	F	No	No	No	No	Euthyroid	Multinodular goiter	No	Papillary	Multi-Focal	Low	T1b	N0	M0	I	Indeterminate	Yes
35	F	No	No	No	No	Euthyroid	Multinodular goiter	Right	Papillary	Multi-Focal	Intermediate	T1b	N1b	M0	I	Actual Incomple	Yes
31	M	Yes	No	No	No	Euthyroid	Multinodular goiter	Right	Papillary	Multi-Focal	Intermediate	T1b	N1b	M0	I	Actual Incomple	Yes
18	F	No	No	No	No	Euthyroid	Single nodular goiter-right	Right	Papillary	Uni-Focal	Intermediate	T1b	N1b	M0	I	Actual Incomple	Yes
62	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Follicular	Multi-Focal	Intermediate	T1b	N0	M0	I	Actual Incomple	Yes
62	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Hurthel cell	Uni-Focal	Low	T2	N0	M0	I	Excellent	No
39	F	No	No	No	No	Euthyroid	Diffuse goiter	No	Hurthel cell	Uni-Focal	Low	T2	N0	M0	I	Indeterminate	No
37	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Hurthel cell	Uni-Focal	Low	T2	N0	M0	I	Excellent	No
26	F	No	Yes	No	No	Subclinical Hypertyroidism	Single nodular goiter-right	No	Hurthel cell	Uni-Focal	Low	T2	N0	M0	I	Excellent	No
24	F	No	No	No	No	Euthyroid	Single nodular goiter-left	No	Hurthel cell	Uni-Focal	Low	T2	N0	M0	I	Indeterminate	No
31	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Follicular	Uni-Focal	Low	T2	N0	M0	I	Excellent	No
57	F	Yes	No	No	No	Euthyroid	Multinodular goiter	No	Follicular	Uni-Focal	Low	T2	N0	M0	I	Excellent	No
28	F	No	No	No	No	Euthyroid	Single nodular goiter-right	No	Follicular	Uni-Focal	Low	T2	N0	M0	I	Excellent	No

KODE PROGRAM

```
import pandas as pd

import numpy as np

from sklearn.preprocessing import LabelEncoder

encoder_gender = LabelEncoder()

encoder_gender.fit(df_tiroid['Gender'])

df_gender = encoder_gender.transform(df_tiroid['Gender'])


encoder_smoking = LabelEncoder()

encoder_smoking.fit(df_tiroid['Smoking'])

df_smoking = encoder_smoking.transform(df_tiroid['Smoking'])


encoder_hxsmoking = LabelEncoder()

encoder_hxsmoking.fit(df_tiroid['Hx Smoking'])

df_hxsmoking = encoder_hxsmoking.transform(df_tiroid['Hx Smoking'])


encoder_hxradiothreapy = LabelEncoder()

encoder_hxradiothreapy.fit(df_tiroid['Hx Radiothreapy'])

df_hxradiothreapy = encoder_hxradiothreapy.transform(df_tiroid['Hx
Radiothreapy'])
```

```
encoder_thyroidfunction = LabelEncoder()

encoder_thyroidfunction.fit(df_tiroid['Thyroid Function'])

df_thyroidfunction = encoder_thyroidfunction.transform(df_tiroid['Thyroid
    Function'])


encoder_physicalexamination = LabelEncoder()

encoder_physicalexamination.fit(df_tiroid['Physical Examination'])

df_physicalexamination =
    encoder_physicalexamination.transform(df_tiroid['Physical Examination'])


encoder_adenopathy = LabelEncoder()

encoder_adenopathy.fit(df_tiroid['Adenopathy'])

df_adenopathy = encoder_adenopathy.transform(df_tiroid['Adenopathy'])


encoder_pathology = LabelEncoder()

encoder_pathology.fit(df_tiroid['Pathology'])

df_pathology = encoder_pathology.transform(df_tiroid['Pathology'])


encoder_focality = LabelEncoder()

encoder_focality.fit(df_tiroid['Focality'])

df_focality = encoder_focality.transform(df_tiroid['Focality'])
```

```
encoder_risk = LabelEncoder()

encoder_risk.fit(df_tiroid['Risk'])

df_risk = encoder_risk.transform(df_tiroid['Risk'])


encoder_T = LabelEncoder()

encoder_T.fit(df_tiroid['T'])

df_T = encoder_T.transform(df_tiroid['T'])


encoder_N = LabelEncoder()

encoder_N.fit(df_tiroid['N'])

df_N = encoder_N.transform(df_tiroid['N'])


encoder_M = LabelEncoder()

encoder_M.fit(df_tiroid['M'])

df_M = encoder_M.transform(df_tiroid['M'])


encoder_stage = LabelEncoder()

encoder_stage.fit(df_tiroid['Stage'])

df_stage = encoder_stage.transform(df_tiroid['Stage'])


encoder_response = LabelEncoder()
```

```

encoder_response.fit(df_tiroid['Response'])

df_response = encoder_response.transform(df_tiroid['Response'])

encoder_recurred = LabelEncoder()

encoder_recurred.fit(df_tiroid['Recurred'])

df_recurred = encoder_recurred.transform(df_tiroid['Recurred'])

df_tiroid = pd.DataFrame({'Gender': df_gender, 'Smoking': df_smoking,
                          'Hx_Smoking': df_hxsmoking, 'Hx_Radiothreapy': df_hxradiothreapy,
                          'Thyroid_Function': df_thyroidfunction,
                          'Physical_Examination': df_physicalexamination,
                          'Adenopathy': df_adenopathy, 'Pathology': df_pathology,
                          'Focality': df_focality, 'Risk': df_risk, 'T': df_T, 'N': df_N, 'M': df_M,
                          'Stage': df_stage, 'Response': df_response, 'Recurred': df_recurred})

y_tiroid = df_tiroid.iloc[:, -1]

x_tiroid = df_tiroid.drop(['Recurred'], axis=1)

from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(x_tiroid, y_tiroid, test_size=0.2,
                                                    random_state=42, shuffle=True)

from sklearn.preprocessing import MinMaxScaler

minmax_norm = MinMaxScaler()

minmax_norm.fit(x_train)

x_train_norm = minmax_norm.transform(x_train)

x_test_norm = minmax_norm.transform(x_test)

```


RIWAYAT HIDUP PENELITI



Nama : Yayan Nasibu
Tempat, Tgl Lahir : Gorontalo, 25 November 2000
Pekerjaan : Mahasiswa
Agama : Islam
Email : yayannasibu@gmail.com
Alamat : Jl jakarta, Kel.Dulalowu,
Kec.Kota Tengah, Kota
Gorontalo.

Riwayat Pendidikan :

1. Sekolah Dasar Negeri 8 Sumalata 2007 s/d 2012
2. SMP Negeri 1 Tibawa 2013 s/d 2015
3. Madrasah Aliyah Muhammadiyah Isimu 2016 s/d 2018
4. Universitas Ichsan Gorontalo 2021 s/d Sekarang



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
LEMBAGA PENELITIAN

Jl. Achmad Nadjamuddin No.17, Kampus Unisan Gorontalo Lt.I Kota Gorontalo 96128
Website: lemlitunisan.ac.id, Email: lembagapenelitian@unisan.ac.id

Nomor : 105/PIP/B.04/LP-UIG/2024
Lampiran : -
Hal : Permohonan Izin Penelitian (Pengambilan Data)

Kepada Yth.,
Universitas Ichsan Gorontalo
di -
Tempat

Yang bertandatangan di bawah ini:

Nama : Dr. Rahmisyari, ST., SE., MM
NIDN : 0929117202
Pangkat Akademik : Lektor Kepala
Jabatan : Ketua Lembaga Penelitian Universitas Ichsan Gorontalo

Meminta kesediaannya untuk memberikan izin pengambilan data dalam rangka penyusunan Proposal/Skripsi, kepada:

Nama : Yayan Nasibu
NIM : T3121033
Fakultas : Ilmu Komputer
Program Studi : Teknik Informatika
Judul Penelitian : Klasifikasi Penyakit Tiroid Menggunakan Algoritma Artificial Neural Network
Lokasi Penelitian : Universitas Ichsan Gorontalo

Demikian surat ini saya sampaikan, atas bantuan dan kerjasamanya diucapkan banyak terima kasih.

Dikeluarkan di Gorontalo

Tanggal 12/10/2024

Ketua Lembaga Penelitian


Dr. Rahmisyari, ST., SE., MM
NIDN: 0929117202



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT KETERANGAN PENELITIAN

Nomor : 040 /FIKOM-UIG/SKP/II/2025

Yang bertanda tangan dibawah ini :

N a m a : Irvan Abraham Salihi, M. Kom
Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini Menerangkan bahwa :

N a m a Mahasiswa : Yayan Nasibu
N I M : T3121033
Program Studi : Teknik Informatika

Bahwa yang bersangkutan benar-benar telah melakukan penelitian tentang “Klasifikasi Penyakit Tiroid Menggunakan Algoritma Artificial Neural Network ” Guna untuk menyelesaikan Studi pada Program Studi Teknik Informatika Fakultas Ilmu Komputer, dan bersangkutan telah menyelesaikan penelitian Tersebut pada TGL 02 Februari 2025 sesuai dengan waktu yang telah di tentukan.

Demikian Surat Keterangan ini dibuat dan digunakan untuk seperlunya.

Gorontalo, 20 Februari 2025
Dekan,

Irvan A. Salihi, S.Kom., M.Kom
NIDN : 0928028101



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
UPT. PERPUSTAKAAN FAKULTAS
SK. MENDIKNAS RI NO. 84/D/0/2001

Jl. Achmad Nadjamuddin No.17 Telp(0435) 829975 Fax. (0435) 829976 Gorontalo

SURAT KETERANGAN BEBAS PUSTAKA

No : 004/Perpustakaan-Fikom/II/2025

Perpustakaan Fakultas Ilmu Komputer (FIKOM) Universitas Ichsan Gorontalo dengan ini menerangkan bahwa :

Nama Anggota : Yayan Nasibu
No. Induk : T3121033
No. Anggota : M20254

Terhitung mulai hari, tanggal : Rabu, 19 Februari 2025, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.

Gorontalo, 19 Februari 2025

**Mengetahui,
Kepala Perpustakaan**

Apriyanto Alhamad, M.Kom

NIDN : 0924048601





KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI

UNIVERSITAS ICHSAN GORONTALO

FAKULTAS ILMU KOMPUTER

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001

Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI

No. 040/FIKOM-UIG/R/II/2025

Yang bertanda tangan di bawah ini :

Nama : Irvan Abraham Salihi, M.Kom
NIDN : 0928028101
Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini menerangkan bahwa :

Nama Mahasiswa : Yayan Nasibu
NIM : T3121033
Program Studi : Teknik Informatika (S1)
Fakultas : Fakultas Ilmu Komputer
Judul Skripsi : Klasifikasi Penyakit Tiroid Menggunakan Algoritma Artificial Neural Network

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi Turnitin untuk judul skripsi di atas diperoleh hasil *Similarity* sebesar 14%, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendeteksian Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.



Irvan Abraham Salihi, M.Kom
NIDN. 0928028101

Gorontalo, 22 Februari 2025
Tim Verifikasi,

Zulfrianto Y. Lamasigi, M.Kom
NIDN. 0914089101

Terlampir :
Hasil Pengecekan Turnitin

Fikom01 Unisan

KLASIFIKASI PENYAKIT TIROID MENGGUNAKAN ALGORITMA ARTIFICIAL NEURAL NETWORK

-  FAKULTAS ILMU KOMPUTER
-  Fak. Ilmu Komputer
-  LL Dikti IX Turnitin Consortium

Document Details

Submission ID

trn:oid::1:3160746450

58 Pages**Submission Date**

Feb 20, 2025, 1:39 PM GMT+7

8,323 Words**Download Date**

Feb 20, 2025, 1:40 PM GMT+7

54,286 Characters**File Name**

SKRIPSI_T3121033_YAYAN_NASIBU_yayannasibugmail.pdf

File Size

1.9 MB

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Top Sources

- 0%  Internet sources
- 9%  Publications
- 12%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Top Sources

0%	Internet sources
9%	Publications
12%	Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Student papers	President University	3%
2	Student papers	Blue Mountain Hotel School	2%
3	Student papers	Universitas Dian Nuswantoro	1%
4	Student papers	Jefferson County School District No. R-1	1%
5	Student papers	Fakultas Teknologi Kebumian dan Energi Universitas Trisakti	<1%
6	Student papers	LL Dikti IX Turnitin Consortium	<1%
7	Student papers	Universitas Negeri Jakarta	<1%
8	Student papers	University of New England	<1%
9	Publication	Apriyanto Alhamad, Azminuddin I. S. Azis, Budy Santoso, Sunarto Taliki. "Prediksi ...	<1%
10	Publication	Sinta Rukiastindari, Luthfia Rohimah, Aprillia Aprillia, Chodidjah Chodidjah, Fara...	<1%
11	Student papers	Universitas Pendidikan Ganesha	<1%

12	Student papers	Walters State Community College	<1%
13	Student papers	Universitas Islam Negeri Sumatera Utara	<1%
14	Publication	Siti Raysyah Raysyah, Veri Arinal, Dadang Iskandar Mulyana. "KLASIFIKASI TING...	<1%
15	Publication	Muhammad Rizky Julianto, Yuma Akbar, Tri Wahyudi. "Analisis Sentimen Respon ...	<1%
16	Publication	Zian Asti Dwiyantri, Cahyo Prianto. "Prediksi Cuaca Kota Jakarta Menggunakan Me...	<1%
17	Publication	M. Reza Noviansyah, Tedy Rismawan, Dwi Marisa Midyanti. "PENERAPAN DATA ...	<1%
18	Student papers	Swinburne University of Technology	<1%
19	Publication	Muhammad Iman Nur Hakim, Joko Siswanto, Aninditya Anggari Nuryono. "Param...	<1%
20	Student papers	Universitas Muria Kudus	<1%



KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS ICHSAN GORONTALO
FAKULTAS ILMU KOMPUTER
SK MENDIKNAS NOMOR 84/D/O/2001

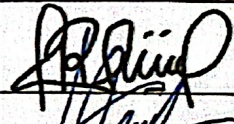
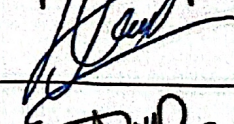
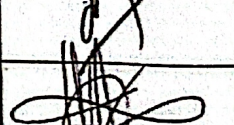
JL. Achmad Nadjamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

Berita Acara Perbaikan/Revisi Ujian Skripsi

Pada hari ini, Senin 24 Februari 2025, Pukul 13.00-14.30 Wita. Telah dilaksanakan Ujian Skripsi mahasiswa Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Nama : Yayan Nasibu
Nim : T3121033
Pembimbing I : Sudirman S. Panna, M.Kom
Pembimbing II : Zulfrianto Y. Lamasigi, M.Kom
Judul Skripsi : Klasifikasi Penyakit Tiroid Menggunakan Algoritma Artificial Neural Network

Oleh Komite Seminar sebagai berikut :

No	Komite Seminar	Status	Tanda Tangan
1	Rezqiwati Ishak, M.Kom	Ketua	
2	Husdi, M.kom	Anggota	
3	Ir. Misrawaty A. Puspa, M.Kom	Anggota	
4	Sudirman S. Panna, M.Kom	Anggota	
5	Zulfrianto Y. Lamasigi, M.Kom	Anggota	