DETEKSI *LUNG CANCER* BERBASIS *GENE EXPRESSION* MENGGUNAKAN *MULTI-LAYER PERCEPTRON* DENGAN *MUTUAL INFORMATION*

SKRIPSI

Oleh :
MUHAMMAD FAYDH MAULA QISTHI
NIM. 220605110143

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026

DETEKSI *LUNG CANCER* BERBASIS *GENE EXPRESSION*
MENGGUNAKAN *MULTI-LAYER PERCEPTRON*
DENGAN *MUTUAL INFORMATION*

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
MUHAMMAD FAYDH MAULA QISTHI
NIM. 220605110143

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026

HALAMAN PERSETUJUAN

DETEKSI *LUNG CANCER* BERBASIS *GENE EXPRESSION* MENGGUNAKAN *MULTI-LAYER PERCEPTRON* DENGAN *MUTUAL INFORMATION*

SKRIPSI

Oleh :
MUHAMMAD FAYDH MAULA QISTHI
NIM. 220605110143

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 24 Juni 2026

Pembimbing I,	Pembimbing II,
Dr Irwan Budi Santoso, M.Kom	Tri Mukti Lestari, M.Kom
NIP. 19770103 201101 1 004	NIP. 19911108 202012 2 005

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Supriyono, M.Kom
NIP. 19841010 201903 1 012

HALAMAN PENGESAHAN

DETEKSI *LUNG CANCER* BERBASIS *GENE EXPRESSION* MENGGUNAKAN *MULTI-LAYER PERCEPTRON* DENGAN *MUTUAL INFORMATION*

SKRIPSI

Oleh :
MUHAMMAD FAYDH MAULA QISTHI
NIM. 220605110143

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: 24 Juni 2025

Susunan Dewan Penguji

Ketua Penguji	: Dr. Agung Teguh Wibowo Almais, M.T. NIP. 19860301 202321 1 016	( )
Anggota Penguji I	: Okta Qomaruddin Aziz, M.Kom NIP. 19911019 201903 1 013	( )
Anggota Penguji II	: Dr Irwan Budi Santoso, M.Kom NIP. 19770103 201101 1 004	( )
Anggota Penguji III	: Tri Mukti Lestari, M.Kom NIP. 19771020 200912 1 001	( )

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang

Supriyono, M. Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

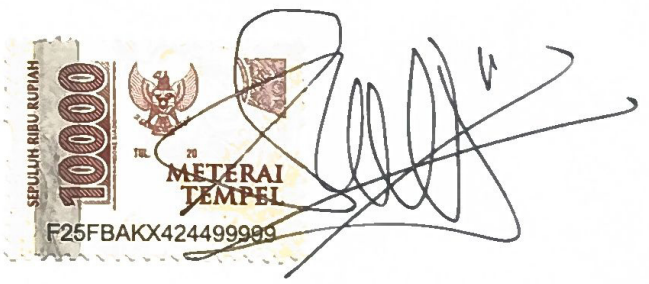
Saya yang bertanda tangan di bawah ini:

Nama	: Muhammad Faydh Maula Qisthi
NIM	: 220605110143
Fakultas / Program Studi	: Sains dan Teknologi / Teknik Informatika
Judul Skripsi	: Deteksi *Lung Cancer* Berbasis *Gene Expression* Menggunakan *Multi-Layer Perceptron* Dengan *Mutual Information*

Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 24 Juni 2026
Yang membuat pernyataan,



Muhammad Faydh Maula Qisthi
NIM.220605110143

MOTTO

"Seorang manusia akan menjadi lebih kuat seiring halangan dan ombak (masalah) yang menerpa menghadangnya"

— Roronoa Zoro (One Piece)

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur, karya ini penulis persembahkan kepada:

Ayah, Ibu, dan adik-adik

atas doa, dukungan, dan kasih sayang tanpa henti yang selalu menjadi sumber kekuatan dan motivasi.

Dosen pembimbing dan para pengajar

atas bimbingan, arahan, serta ilmu yang sangat berarti bagi penulis.

Teman-teman Infinity Teknik Informatika 2022

yang telah menjadi bagian dari perjalanan ini melalui kebersamaan dan semangat yang tidak pernah padam.

Dan akhirnya, untuk diri sendiri

yang terus berusaha, melangkah, dan bertahan hingga mencapai titik ini.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahirabbil'alamin, segala puji dan syukur senantiasa penulis haturkan kepada Allah SWT. yang telah melimpahkan rahmat serta bimbingan-Nya, sehingga penulis berhasil menuntaskan skripsi berjudul " DETEKSI *LUNG CANCER* BERBASIS *GENE EXPRESSION* MENGGUNAKAN *MULTI-LAYER PERCEPTRON* DENGAN *MUTUAL INFORMATION*" dengan baik. Shalawat dan salam penulis curahkan kepada Rasulullah Muhammad SAW., panutan agung umat manusia yang telah menuntun kita dari masa kebodohan menuju cahaya kebenaran, agama Islam, serta ilmu pengetahuan yang hingga kini masih kita rasakan manfaatnya.

Skripsi ini disusun sebagai salah satu persyaratan untuk meraih gelar Sarjana Komputer di Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, UIN Maulana Malik Ibrahim Malang. Selama proses pengerjaannya, penulis senantiasa mendapat bantuan, motivasi, dan dukungan doa dari berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si., selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Dr. H. Agus Mulyono, S.Pd., M.Kes, selaku Dekan Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang.

3. Supriyono, M. Kom., selaku Ketua Program Studi Teknik Informatika UIN Maulana Malik Ibrahim Malang.
4. Dr Irwan Budi Santoso, M.Kom., selaku Dosen Pembimbing I, yang dengan kesabaran dan dedikasinya membimbing penulis selama penyusunan skripsi.
5. Tri Mukti Lestari, M.Kom., selaku Dosen Pembimbing II, yang senantiasa memberikan arahan dan masukan yang sangat berharga dalam pengerjaan skripsi ini.
6. Dr. Agung Teguh Wibowo Almais, M.T., dan Okta Qomaruddin Aziz, M.Kom., selaku Dosen Penguji yang telah memberikan kritik, saran, dan bimbingan untuk menyempurnakan skripsi ini.
7. Nia Faricha, S.Si., admin Program Studi Teknik Informatika, yang membantu dalam urusan administrasi dan selalu mengingatkan tentang kelengkapan berkas.
8. Ayah Muhammad Abdul Kholiq dan Ibu Arifatul Hidayah, S.Ag., orang tua tercinta, serta Muhammad Rafif Rassendriya ‘Azmi dan Nadine Ghaydatul Zahira, adik-adik saya, yang selalu menjadi pilar kekuatan dengan doa, cinta, dan dukungan tanpa henti.
9. Putri Lestari, yang dengan ketulusan, kesabaran, dan dukungannya dari jauh selalu menjadi sumber motivasi serta kekuatan bagi penulis dalam menyelesaikan tugas akhir ini.
10. Rekan seperjuangan “2RU”, yang beranggotakan Rheza, Sada, Izzan, Fiki, Lana. Karya kecil ini kupersembahkan untuk kalian yang selalu menjadi

sumber semangatku. Terima kasih sudah mau berbagi tawa dan air mata selama ini.

11. Teman solid "PDI" yang beranggotakan Ridiey dan Syafri. Yang membersamai baik ketika berada dikala senang maupun sedih.
12. Seluruh keluarga, teman, sahabat, dan kerabat penulis yang tidak dapat penulis sebutkan satu per satu, yang turut memberikan bantuan, semangat, dukungan, dan doa untuk penulis.

Dan terakhir tak lupa terimakasih kepada diri sendiri, yang dengan tekad kuat mampu melewati setiap tantangan, rintangan, dan kesulitan selama perjalanan ini, membuktikan bahwa kerja keras dan usaha tidak pernah sia-sia. Penulis menyadari bahwa penelitian dalam tugas skripsi ini masih jauh dari kata sempurna dan memiliki berbagai keterbatasan. Oleh karena itu, dengan penuh kerendahan hati, penulis membuka diri untuk menerima kritik dan saran yang membangun dari para pembaca guna menjadi bahan evaluasi dan pengembangan di masa mendatang. Penelitian ini juga memiliki potensi untuk dilanjutkan dan dikembangkan lebih jauh dalam penelitian berikutnya, sehingga dapat melengkapi kekurangan yang ada. Penulis berharap, karya ini tidak hanya memberikan manfaat bagi pembaca, tetapi juga dapat berkontribusi positif bagi masyarakat secara luas.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Malang, 24 Juni 2026

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

ABSTRAK

Qisthi, Muhammad Faydh Maula. 2026. **Deteksi *Lung Cancer* Berbasis *Gene expression* Menggunakan *MultiLayer Perceptron* Dengan *Mutual Information* .** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr Irwan Budi Santoso, M.Kom. (II) Tri Mukti Lestari, M.Kom.

Kata kunci: Deteksi Kanker Paru, Ekspresi Gen, *Multi-Layer Perceptron*, Mutual Information, TCGA-LUSC.

Kanker paru-paru merupakan salah satu penyebab kematian tertinggi di dunia, dengan tingkat kelangsungan hidup yang rendah akibat keterlambatan diagnosis. Pendekatan diagnostik berbasis data ekspresi gen menawarkan potensi deteksi dini pada tingkat molekuler, namun karakteristik data yang berdimensi sangat tinggi dengan jumlah sampel terbatas menyebabkan metode pembelajaran mesin konvensional seperti SVM dan *Random Forest* rentan mengalami penurunan performa, sementara metode *deep learning* yang lebih kompleks memerlukan sumber daya komputasi besar. Penelitian ini bertujuan untuk mengetahui dan menganalisis performa metode *Multi-Layer Perceptron* (MLP) dalam mendeteksi data ekspresi gen guna membedakan jaringan paru normal dan jaringan yang terindikasi kanker. Dataset yang digunakan adalah TCGA-LUSC yang terdiri dari 551 sampel dengan 56.907 fitur gen hasil RNA-*sequencing* yang telah dinormalisasi menggunakan metode TPM. Tahap prapemrosesan dilakukan melalui standarisasi data, sedangkan seleksi fitur dilakukan menggunakan *Mutual Information* untuk memilih gen-gen yang paling relevan. Selanjutnya, berbagai konfigurasi arsitektur MLP diuji dengan variasi jumlah *hidden layer*, neuron, dan fitur input. Hasil pengujian melalui skema 5-*Fold Cross validation* menunjukkan bahwa model terbaik, dengan arsitektur tiga *hidden layer* berkonfigurasi neuron (64, 32, 16) dan 100 fitur input, mampu mencapai akurasi rata-rata sebesar 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Hasil tersebut mengindikasikan bahwa kombinasi MLP dan *Mutual Information* efektif serta efisien dalam mengklasifikasikan data ekspresi gen berdimensi tinggi, sehingga berpotensi diterapkan sebagai alat bantu pendukung keputusan klinis dalam deteksi dini kanker paru.

ABSTRACT

Qisthi, Muhammad Faydh Maula. 2026. ***Gene Expression*-Based *Lung Cancer* Detection Using a *Multi*-Layer *Perceptron* with Mutual Information**. Thesis. Department of Computer Science, Faculty of Science and Technology, Maulana Malik Ibrahim State Islamic University, Malang. Advisors: (I) Dr. Irwan Budi Santoso, M.Kom. (II) Tri Mukti Lestari, M.Kom.

Lung cancer is one of the leading causes of death worldwide, with low survival rates due to delayed diagnosis. Diagnostic approaches based on *gene expression* data offer the potential for early detection at the molecular level; however, the highly dimensional nature of the data combined with a limited number of samples causes conventional *machine learning* methods such as SVM and *Random Forest* to be prone to performance degradation, while more complex *deep learning* methods require significant computational resources. This study aims to investigate and analyze the performance of the *Multi*-Layer *Perceptron* (MLP) method in detecting *gene expression* data to distinguish between normal *lung* tissue and tissue suspected of containing *cancer*. The dataset used is TCGA-LUSC, consisting of 551 samples with 56,907 *gene feature* s derived from RNA-*sequencing* that have been normalized using the TPM method. The *preprocessing* stage involved data standardization, while *feature Selection* was performed using *Mutual Information* to identify the most relevant genes. Next, various MLP architecture configurations were tested with different numbers of hidden layers, neurons, and input *feature* s. The results of testing using a 5-fold cross-validation scheme showed that the best model—with an architecture of three hidden layers configured with (64, 32, 16) neurons and 100 input *feature* s—achieved an average accuracy of 99,64%, a precision of 99,65%, a recall of 99,64%, and an F1-score of 99,64%. These results indicate that the combination of MLP and *Mutual Information* is both effective and efficient in classifying high-dimensional *gene expression* data, making it a potential tool for supporting clinical decision-making in the early detection of *lung cancer*.

Keyword: *Gene Expression*, *Lung Cancer* Detection, *Multi*-Layer *Perceptron*, Mutual Information, TCGA-LUSC.

مستخلص البحث

قيثي، محمد فايد مولا. 2026. **الكشف عن سرطان الرئة استنادًا إلى التعبير الجيني باستخدام شبكة عصبية متعددة الطبقات مع المعلومات المتبادلة. أطروحة تخرج**. قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، مالانج. المشرفان (I) :د. إيروان بودي سانتوسو، ماجستير في علوم الحاسوب (II) .تري موكتي ليستاري، ماجستير في علوم الحاسوب.

الكلمات المفتاحية :الكشف عن سرطان الرئة، التعبير الجيني، الشبكة العصبية المتعددة الطبقات، المعلومات المتبادلة، TCGA-LUSC.

يُعد سرطان الرئة أحد أكثر أسباب الوفاة شيوعًا في العالم، مع انخفاض معدلات البقاء على قيد الحياة نتيجة التأخر في التشخيص. تقدم المناهج التشخيصية القائمة على بيانات التعبير الجيني إمكانية الكشف المبكر على المستوى الجزيئي، إلا أن خصائص البيانات ذات الأبعاد العالية جدًّا مع عدد العينات المحدود تجعل طرق التعلم الآلي التقليدية مثل *Random Forest* و *SVM* عرضة لانخفاض الأداء، في حين تتطلب طرق التعلم العميق الأكثر تعقيدًا موارد حاسوبية ضخمة. تهدف هذه الدراسة إلى تحديد وتحليل أداء طريقة *MLP (Multi-Layer Perceptron)* في تحليل بيانات التعبير الجيني للتمييز بين أنسجة الرئة السليمة والأنسجة التي تشير إلى وجود سرطان. مجموعة البيانات المستخدمة هي TCGA-LUSC التي تتألف من 551 عينة تحتوي على 56,907 سمة جينية ناتجة عن تسلسل الحمض النووي الريبي (*RNA-sequencing*) والتي تمت معايرتها باستخدام طريقة TPM. تم تنفيذ مرحلة المعالجة المسبقة من خلال توحيد البيانات، في حين تم اختيار السمات باستخدام «المعلومات المتبادلة» (*Mutual Information*) مع تغيير عدد الطبقات المخفية، والخلايا العصبية، وسمات الإدخال. أظهرت نتائج الاختبار من خلال مخطط التحقق المتقاطع الخماسي (*5-Fold Cross validation*) أن النموذج الأفضل، الذي يتكون من ثلاث طبقات مخفية بتكوين الخلايا العصبية (64، 32، 16) و100 سمة إدخال، تمكن من تحقيق دقة متوسطة بلغت 99,64%، ودقة بنسبة 99,65%، ومعدل استرجاع بنسبة 99,64%، ودرجة F1 بنسبة 99,64%. تشير هذه النتائج إلى أن الجمع بين شبكة MLP فعال وفعال (*Mutual Information*) والمعلومات المتبادلة في تصنيف بيانات التعبير الجيني عالية الأبعاد، مما يجعله قابلاً للتطبيق كأداة مساعدة لدعم القرار السريري في الكشف المبكر عن سرطان الرئة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker paru-paru merupakan salah satu penyebab kematian tertinggi di dunia dan menjadi beban kesehatan global yang signifikan. Berdasarkan laporan *World Health Organization* (WHO), kanker paru menyumbang lebih dari 18% kematian akibat kanker pada tahun 2024, dengan tingkat kelangsungan hidup lima tahun yang masih rendah terutama karena keterlambatan diagnosis (Lin & Bao, 2024). Kanker paru sering kali baru terdeteksi pada stadium lanjut, ketika gejala klinis mulai muncul dan penyebaran telah terjadi ke organ lain. Kondisi ini semakin diperburuk oleh kenyataan bahwa sebagian besar pasien baru terdiagnosis ketika penyakit sudah mencapai stadium lanjut, sehingga pilihan terapi menjadi terbatas dan tingkat kelangsungan hidup menurun drastis.

Berbagai penelitian menunjukkan bahwa peningkatan kesadaran masyarakat terhadap faktor risiko, seperti merokok, paparan polusi udara, serta faktor lingkungan dan genetik, merupakan langkah awal penting dalam mengurangi angka kematian akibat kanker paru (Bychowski & Sobiczewski, 2023; Deo et al., 2022). Selain itu, prediksi global memperkirakan bahwa pada tahun 2050 jumlah kasus kanker paru dapat mencapai 3,8 juta, menandakan perlunya strategi yang lebih kuat dalam pencegahan, deteksi dini, dan kebijakan kesehatan publik yang terpadu (Sharma, 2022). Dengan demikian, kolaborasi lintas sektor antara pemerintah, lembaga kesehatan, dan komunitas penelitian menjadi sangat penting

untuk memperkuat sistem kesehatan dan memperluas akses terhadap diagnosis serta perawatan kanker paru yang efektif.

Dalam konteks keilmuan Islam, penelitian ini sejalan dengan perintah Allah SWT agar manusia menggunakan akal dan pikirannya untuk meneliti ciptaan-Nya. Al-Qur'an menegaskan dalam QS. Al-Mulk [67]:3–4,

ٱلَّذِي خَلَقَ سَبۡعَ سَمَٰوَٰتٖ طِبَاقٗاۖ مَّا تَرَىٰ فِي خَلۡقِ ٱلرَّحۡمَٰنِ مِن تَفَٰوُتٖۖ فَٱرۡجِعِ ٱلۡبَصَرَ هَلۡ تَرَىٰ مِن فُطُورٖ ٣
ثُمَّ ٱرۡجِعِ ٱلۡبَصَرَ كَرَّتَيۡنِ يَنقَلِبۡ إِلَيۡكَ ٱلۡبَصَرُ خَاسِئٗا وَهُوَ حَسِيرٞ ٤

"*Yang telah menciptakan tujuh langit berlapis-lapis. Tidaklah kamu lihat sesuatu yang tidak seimbang pada ciptaan Tuhan Yang Maha Pemurah; maka lihatlah sekali lagi, adakah kamu lihat sesuatu yang cacat? Kemudian pandanglah sekali lagi niscaya penglihatanmu akan kembali kepadamu dengan tidak menemukan sesuatu cacat dan penglihatanmu itu pun dalam keadaan letih."* (*QS. Al-Mulk:3-4*)

Surat Al-Mulk ayat 3–4 menjelaskan tentang kesempurnaan ciptaan Allah SWT serta perintah kepada manusia untuk memperhatikan dan merenungkan penciptaan alam semesta. Menurut tafsir Ibnu Katsir, ayat tersebut menegaskan bahwa Allah menciptakan langit secara berlapis-lapis dengan susunan yang sempurna tanpa adanya ketidakseimbangan, cacat, atau kerusakan (Al Shaykh & Abdul Ghoffar, 2006). Oleh karena itu manusia diperintahkan untuk melihat dan meneliti kembali ciptaan tersebut secara berulang-ulang untuk memastikan bahwa tidak terdapat kekurangan dalam penciptaan-Nya. Perintah untuk "melihat kembali" menunjukkan dorongan bagi manusia untuk melakukan pengamatan secara mendalam terhadap alam semesta dan segala sistem yang ada di dalamnya.

Dalam perspektif ilmu pengetahuan, dorongan untuk mengamati dan meneliti ciptaan Allah dapat diwujudkan melalui kegiatan penelitian ilmiah. Penelitian dalam bidang biologi dan bioinformatika, seperti analisis ekspresi gen

untuk mendeteksi penyakit kanker paru, merupakan salah satu bentuk upaya manusia dalam memahami keteraturan dan kompleksitas sistem biologis yang diciptakan oleh Allah SWT. Dengan demikian, kegiatan penelitian ilmiah dapat menjadi sarana untuk mengenal lebih jauh kebesaran dan kesempurnaan ciptaan-Nya.

Sejalan dengan meningkatnya kebutuhan akan deteksi dini kanker paru, berbagai penelitian telah memanfaatkan data ekspresi gen sebagai pendekatan diagnostik berbasis molekuler. Penelitian yang dilakukan oleh (Mohamed et al., 2023) menunjukkan bahwa profil ekspresi gen mampu membedakan jaringan paru normal dan jaringan kanker melalui analisis pola aktivitas genetik yang berubah akibat proses karsinogenesis. Sementara itu, (Tabassum et al., 2024) memanfaatkan teknologi *microarray* dan RNA *sequencing* untuk mengukur ribuan gen secara simultan, sehingga menghasilkan representasi biologis sel yang lebih komprehensif. Selain itu, menegaskan bahwa analisis ekspresi gen berperan penting dalam identifikasi biomarker dan subtipe kanker paru, yang dapat mendukung pemilihan terapi yang lebih tepat sasaran. Meskipun demikian, (Yuvaraj & Maheswari, 2023) menjelaskan bahwa data ekspresi gen memiliki dimensi fitur yang sangat tinggi dengan jumlah sampel yang cenderung terbatas, sehingga meningkatkan risiko *overfitting* dan menurunkan kemampuan generalisasi model. Pendekatan pembelajaran mesin tradisional, seperti SVM dan *Random Forest*, juga dilaporkan mengalami penurunan performa ketika dihadapkan pada kompleksitas korelasi gen yang tinggi (Nassif et al., 2025a). Sementara itu, metode *deep learning* seperti CNN, *Autoencoder*, dan RNN menawarkan akurasi yang lebih baik, namun

memerlukan arsitektur yang kompleks, waktu pelatihan yang panjang, dan sumber daya komputasi yang besar (Sun et al., 2024; Chen et al., 2022). Bahkan, (Yuvaraj & Maheswari, 2023) yang menggabungkan MLP dengan algoritma optimasi IMPSO untuk meningkatkan akurasi klasifikasi kanker paru tetap menghadapi tantangan pada efisiensi komputasi dan kebutuhan GPU berperforma tinggi. Kondisi ini menunjukkan bahwa masih diperlukan model klasifikasi yang lebih ringan, efisien, namun tetap akurat, sehingga dapat diterapkan secara praktis pada analisis data ekspresi gen berdimensi tinggi.

Untuk mengatasi permasalahan tersebut, *Multilayer Perceptron* (MLP) menjadi salah satu pendekatan yang relevan dan efisien dalam klasifikasi berbasis ekspresi gen. MLP merupakan arsitektur jaringan saraf tiruan yang mampu mempelajari hubungan non-linear antar fitur, sehingga cocok untuk menganalisis pola kompleks pada data genetik berdimensi tinggi. Struktur MLP yang tersusun dari lapisan input, beberapa lapisan tersembunyi, dan lapisan *output* memungkinkan model melakukan penyesuaian bobot secara bertahap melalui proses *backpropagation*, sehingga representasi fitur yang lebih informatif dapat terbentuk (Jiang, 2024). Beberapa penelitian terkini menunjukkan bahwa MLP mampu memberikan performa klasifikasi yang kompetitif pada data biomedis, sekaligus memiliki kompleksitas komputasi yang lebih rendah dibandingkan arsitektur *deep learning* yang lebih kompleks seperti CNN atau RNN (Alharbi & Vakanski, 2023; Tabassum et al., 2024). Dengan demikian, penggunaan MLP diharapkan dapat menghasilkan model yang lebih ringan, cepat, dan tetap akurat, sehingga lebih potensial untuk mendukung proses deteksi dini kanker paru berbasis ekspresi gen.

1.2 Rumusan Masalah

Bagaimana performa metode *Multilayer Perceptron* (MLP) dalam mendeteksi data ekspresi gen untuk membedakan jaringan paru normal dan jaringan yang terindikasi kanker ?

1.3 Batasan Masalah

1. Sumber Data: Data yang digunakan sepenuhnya bersumber dari dataset sekunder TCGA-LUSC (*Lung Squamous Cell Carcinoma*) yang diperoleh melalui portal GDC atau Kaggle (noepinefrin, 2024).
2. Reduksi Dimensi (Seleksi Fitur): Mengingat tingginya dimensi data yang mencapai 56.907 gen, tidak semua gen akan digunakan sebagai input. Penelitian ini dibatasi pada penerapan metode *feature Selection* guna menyaring gen-gen yang paling relevan.
3. Variasi Pengujian Model: Pengujian model dibatasi pada evaluasi performa *Multi-Layer Perceptron* (MLP) terhadap perbandingan berbagai jumlah fitur input terbaik. Tujuannya adalah untuk menemukan jumlah fitur yang paling optimal untuk proses klasifikasi.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui dan menganalisis performa metode *Multilayer Perceptron* (MLP) dalam mendeteksi data ekspresi gen sehingga mampu membedakan jaringan paru normal dan jaringan yang terindikasi kanker.

1.5 Manfaat Penelitian

1. Memberikan peluang deteksi dini kanker paru pada tingkat molekuler sehingga meningkatkan potensi kesembuhan dan harapan hidup pasien.
2. Menjadi alat bantu bagi tenaga medis dalam pengambilan keputusan klinis melalui analisis data genomik yang cepat, objektif, dan akurat.
3. Membantu rumah sakit meningkatkan efisiensi pelayanan dengan sistem deteksi otomatis yang akurat serta mengoptimalkan penggunaan sumber daya medis.

BAB II

STUDI PUSTAKA

2.1 Klasifikasi Kanker Paru Berbasis Ekspresi Gen dan *Mechine learning*

Penelitian mengenai klasifikasi kanker menggunakan data *gene expression* telah banyak dilakukan dengan memanfaatkan berbagai teknik *machine learning* dan *deep learning*. Salah satu penelitian dilakukan oleh Akter et al., (2025) yang mengusulkan pendekatan *machine learning* untuk mengidentifikasi gen signifikan serta mengklasifikasikan berbagai jenis kanker menggunakan data RNA *sequencing*. Dataset yang digunakan berasal dari Gene expression Cancer RNA-Seq dari TCGA yang terdiri dari 801 sampel dengan lebih dari 20.000 fitur gen. Dalam penelitian tersebut dilakukan proses seleksi fitur menggunakan metode *Lasso Regression* dan *Ridge Regression* untuk mengurangi dimensi data dan mengidentifikasi gen yang paling berpengaruh. Selanjutnya beberapa algoritma *machine learning* seperti *Support Vector Machine* (SVM), *Random Forest*, *K-Nearest Neighbor, Decision Tree*, dan *Naïve Bayes* digunakan untuk melakukan klasifikasi. Hasil penelitian menunjukkan bahwa algoritma SVM memberikan performa terbaik dengan akurasi mencapai 99,87% pada proses validasi menggunakan 5-*fold cross validation*.

Penelitian lain dilakukan oleh Sheet et al., (2024) yang mengusulkan model *deep learning* untuk mengenali gen yang berperan dalam perkembangan kanker menggunakan data *gene expression*. Penelitian tersebut menggunakan model *Multilayer Perceptron* yang dikombinasikan dengan *Stacked Denoising*

Autoencoder (MLP-SDAE) untuk mengekstraksi fitur penting dari data genomik. Model *autoencoder* digunakan untuk mempelajari representasi fitur laten dari data gen yang berdimensi tinggi sebelum dilakukan proses klasifikasi menggunakan MLP. Pendekatan ini bertujuan untuk meningkatkan kemampuan model dalam mengidentifikasi gen yang berperan dalam proses mediasi kanker. Hasil penelitian menunjukkan bahwa kombinasi metode *autoencoder* dan neural network mampu meningkatkan akurasi klasifikasi serta membantu dalam proses identifikasi gen penting yang berkaitan dengan perkembangan kanker.

Penelitian yang dilakukan oleh Tabassum et al. (2024) meneliti klasifikasi kanker berbasis *gene expression* dengan mengusulkan metode MI-Bagging, yaitu kombinasi *Mutual Information* (MI) sebagai teknik seleksi fitur dan *bagging ensemble* dengan *Multilayer Perceptron* (MLP) sebagai base learner, karena data *gene expression* memiliki jumlah fitur yang sangat tinggi tetapi jumlah sampel relatif kecil sehingga berisiko menimbulkan *overfitting* dan menurunkan akurasi model; penelitian ini menggunakan lima dataset benchmark yang mencakup leukemia (DS1), *brain cancer* (DS2), *Multi-cancer* 5 kelas berisi BRCA, KIRC, LUAD, PRAD, dan COAD (DS3), *Multi-cancer* 11 kelas (DS4), serta *ovarian cancer* (DS5) dengan jumlah fitur antara 4.656–16.383 gen dan jumlah sampel antara 72–801, kemudian dilakukan tahap praproses berupa pengecekan *missing value*, SMOTE untuk menyeimbangkan kelas, label encoding, pembagian data latih dan uji 80:20, serta normalisasi data ke rentang 0–1; selanjutnya fitur diperingkat menggunakan MI dan dipilih sejumlah fitur terbaik melalui *trial and error*, lalu diklasifikasikan menggunakan 10 model MLP hasil bootstrap sampling, di mana

setiap MLP memiliki 3 *hidden layer* dengan 50 neuron per *layer*, fungsi aktivasi ReLU, *solver* Adam, dan *learning rate* 0,001; hasil penelitian menunjukkan bahwa metode yang diusulkan mampu mencapai akurasi 100% pada DS1, DS2, dan DS3, 98,48% pada DS4, serta 95,83% pada DS5, dengan jumlah fitur terbaik masing-masing 100, 500, 500, 1500, dan 2000 gen, serta secara umum mengungguli beberapa metode sebelumnya seperti mRMR-ANN, PCA-XGBoost, PCA-ANN, PCA-*Random Forest*, *Genetic Programming*, SVM, NN, dan CNN; namun demikian, penelitian ini tidak berfokus khusus pada *lung cancer* LUSC dan tidak menggunakan MLP tunggal, melainkan klasifikasi berbagai jenis kanker dengan pendekatan *ensemble bagging* berbasis MLP, sehingga relevan dengan penelitian ini pada aspek penggunaan *gene expression*, MI, dan MLP, tetapi berbeda pada fokus dataset dan arsitektur model yang digunakan.

Penelitian lainnya dilakukan oleh Lima et al., (2025) yang mengembangkan pendekatan *machine learning* yang dikombinasikan dengan *Explainable Artificial Intelligence* (XAI) untuk menemukan biomarker yang berkaitan dengan stadium kanker paru jenis *Lung Squamous Cell Carcinoma* (LUSC). Dataset yang digunakan berasal dari TCGA-LUSC yang terdiri dari 493 pasien. Dalam penelitian ini dilakukan proses seleksi fitur menggunakan metode *Recursive Feature Elimination* (RFE) untuk mengurangi jumlah gen yang digunakan sebagai fitur input. Selain itu, teknik penyeimbangan data seperti SMOTE dan ADASYN digunakan untuk mengatasi permasalahan ketidakseimbangan kelas pada dataset. Beberapa algoritma *machine learning* seperti *K-Nearest Neighbor*, *Support Vector Machine*, *Logistic Regression*, *Random Forest*, *Multilayer Perceptron*, XGBoost,

dan CatBoost digunakan dalam proses klasifikasi. Hasil penelitian menunjukkan bahwa *Random Forest* menghasilkan performa terbaik dengan akurasi sekitar 0,91 serta nilai AUC yang tinggi.

Penelitian yang dilakukan oleh Shah et al., (2026) mengusulkan kerangka kerja *hybrid feature extraction* untuk klasifikasi kanker paru menggunakan data *gene expression*. Dalam penelitian tersebut digunakan kombinasi *Principal Component Analysis* (PCA) dan *Mutual Information* (MI) untuk mengurangi dimensi data dan memilih fitur yang paling relevan terhadap klasifikasi kanker. PCA digunakan untuk menangkap variasi utama dari data yang berdimensi tinggi, sedangkan *Mutual Information* digunakan untuk mengukur tingkat relevansi fitur terhadap kelas target. Setelah proses ekstraksi fitur dilakukan, data kemudian digunakan sebagai input untuk model *deep learning berupa Convolutional Neural Network* (CNN). Hasil penelitian menunjukkan bahwa pendekatan *hybrid* PCA-MI mampu meningkatkan performa klasifikasi kanker paru dengan akurasi mencapai sekitar 98%.

Penelitian lain yang berkaitan dengan klasifikasi kanker paru dilakukan oleh Nassif et al., (2025) yang mengusulkan model *deep learning* berbasis *Convolutional Neural Network* (CNN) untuk mengklasifikasikan tingkat keparahan kanker paru menggunakan data *gene expression*. Dataset yang digunakan mencakup dua jenis kanker paru yaitu *Lung Adenocarcinoma* (LUAD) dan *Lung Squamous Cell Carcinoma* (LUSC). Dalam penelitian ini dilakukan proses seleksi fitur menggunakan metode *F-test* untuk menentukan gen yang paling relevan terhadap klasifikasi penyakit. Model CNN kemudian dilatih menggunakan data

yang telah melalui proses *preprocessing* dan *feature selection*. Hasil penelitian menunjukkan bahwa model CNN yang diusulkan mampu mencapai tingkat akurasi sebesar 93,94% untuk klasifikasi LUAD dan 88,42% untuk klasifikasi LUSC.

Penelitian lain dilakukan oleh Gu yang mengembangkan model *machine learning* untuk memprediksi stadium kanker paru menggunakan data ekspresi gen tumor. Dataset yang digunakan berasal dari TCGA dan terdiri dari 992 sampel dengan hampir 20,000 fitur gen. Dalam penelitian ini dilakukan proses normalisasi data menggunakan transformasi Log2 untuk meningkatkan stabilitas model. Selanjutnya digunakan metode *Wilcoxon Rank Sum Test* untuk mengidentifikasi gen yang memiliki perbedaan signifikan antara stadium awal dan stadium lanjut kanker paru. Setelah proses seleksi fitur awal dilakukan, metode *Incremental Feature Selection* (IFS) digunakan untuk menentukan jumlah fitur optimal yang digunakan dalam proses klasifikasi. Algoritma *machine learning* seperti *Support Vector Machine*, *Random Forest*, dan XGBoost digunakan untuk membangun model klasifikasi. Hasil penelitian menunjukkan bahwa model XGBoost memberikan performa terbaik dalam memprediksi stadium kanker paru berdasarkan data *gene expression*.

Penelitian lain yang berkaitan dengan analisis data kanker menggunakan pendekatan *machine learning* juga menunjukkan bahwa penggunaan teknik seleksi fitur dan metode klasifikasi yang tepat sangat penting dalam menangani data biologis yang memiliki dimensi tinggi. Penelitian tersebut memanfaatkan berbagai teknik *preprocessing*, *feature selection*, dan algoritma klasifikasi untuk meningkatkan performa model dalam menganalisis data genomik. Hasil penelitian

menunjukkan bahwa kombinasi teknik *preprocessing*, seleksi fitur, serta algoritma *machine learning* yang tepat dapat meningkatkan akurasi klasifikasi penyakit berbasis data biologis dan membantu proses diagnosis secara lebih efektif. Tabel 2.1 berikut merangkum perbandingan metodologis dan hasil dari penelitian-penelitian utama tersebut.

Tabel 2. 1 Tabel Penelitian Terkait

No	Penulis	Judul Penlitian	Metode	Hasil	Perbedaan Penelitian
1	Akter et al. (2025)	*Machine learning Approach to Identify Significant Genes and Classify Cancer Types from RNA-seq Data*	*Lasso Regression, Ridge Regression*, SVM, KNN, *Random Forest*, Decision Tree	SVM menghasilkan akurasi tertinggi sebesar 99,87% pada proses klasifikasi jenis kanker	Penelitian ini berfokus pada penggunaan MLP untuk klasifikasi kanker paru
2	Sheet et al.(2024)	*Recognition of Cancer Mediating Genes Using MLP-SDAE Model*	*Multilayer Perceptron* (MLP), *Stacked Denoising Autoencoder* (SDAE)	Model MLP-SDAE mampu mengidentifikasi gen yang berperan dalam perkembangan kanker dengan performa klasifikasi yang baik	Penelitian ini berfokus pada ekstraksi fitur menggunakan MI tanpa *Autoencoder* pada model MLP.
3	Tabassum et al. (2024)	*Cancer Classification from Gene expression Using Ensemble Learning*	Seleksi fitur *Mutual Information* (MI) + *Bagging* dengan MLP sebagai *base learner*	Akurasi tertinggi 100% pada DS1, DS2, DS3; 98,48% pada DS4; 95,83% pada DS5	Penelitian ini berfokus pada penggunaan model MLP tunggal tanpa metode *ensemble learning*.
4	Lima et al. (2025)	*Explainable AI and Multiclassifiers for Staging Biomarker Discovery in Lung Squamous Cell Carcinoma*	RFE *Feature Selection,* SMOTE, ADASYN, *Random Forest*, SVM, MLP, XGBoost	*Random Forest* memberikan performa terbaik dengan akurasi sekitar 0,91 dan mampu mengidentifikasi biomarker kanker paru	Penelitian ini berfokus pada penggunaan MI untuk klasifikasi biner keberadaan kanker paru.
5	Shah et al. (2026)	*A Hybrid Feature Extraction Framework Combining PCA and Mutual Information for Gene expression Based Lung*	PCA, *Mutual Information,* CNN	CNN mencapai akurasi sekitar 98% dalam klasifikasi kanker paru	Penelitian ini berfokus pada penggunaan MI tanpa PCA dengan model MLP untuk klasifikasi.

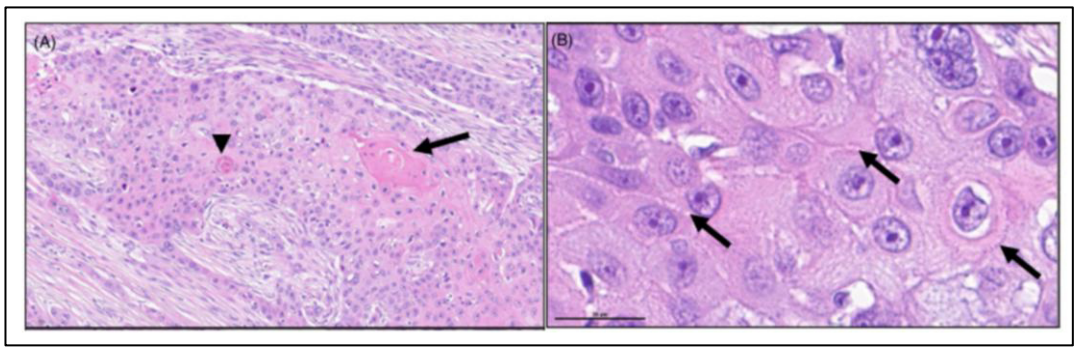
		Cancer Classification			
6	Nassif et al. (2025)	*Classification of Lung Cancer Severity Using Gene expression Data Based on Deep learning*	*Convolutional Neural Network* (CNN), *Feature Selection (F-test)*	CNN mencapai akurasi 93,94% untuk LUAD dan 88.42% untuk LUSC	Penelitian ini berfokus pada seleksi fitur menggunakan MI dan klasifikasi menggunakan MLP.
7	Gu (2024)	*Prediction of Lung Cancer Stage Using Tumor Gene expression Data*	*Wilcoxon Rank Sum Test, Incremental Feature Selection* (IFS), SVM, *Random Forest*, XGBoost	XGBoost memberikan performa terbaik dalam memprediksi stadium kanker paru	Penelitian ini berfokus pada deteksi awal (klasifikasi biner) menggunakan MLP, bukan prediksi stadium.
8	Alzoubi et al. (2023)	*Machine learning Based Cancer Detection Using Gene expression Data*	*Machine learning*, *Feature Selection, Classification Algorithms*	Metode *machine learning* mampu meningkatkan akurasi klasifikasi kanker berbasis data biologis	Penelitian ini berfokus pada optimalisasi MLP dan MI khusus untuk deteksi kanker paru.

2.2 Kanker Paru Sel Skuamosa (LUSC)

Kanker paru merupakan salah satu penyebab utama kematian akibat kanker di seluruh dunia. Berdasarkan laporan Global Cancer Statistics, kanker paru memiliki angka insidensi dan mortalitas yang tinggi dibandingkan jenis kanker lainnya. Salah satu subtipe utama dari kanker paru adalah *Lung Squamous Cell Carcinoma* (LUSC) yang termasuk dalam kelompok *Non-Small Cell Lung Cancer* (NSCLC). Subtipe ini menyumbang sekitar 20–30% dari seluruh kasus kanker paru dan umumnya berkaitan erat dengan kebiasaan merokok jangka panjang (Luo et al., 2025).

Secara histopatologis, LUSC berasal dari sel epitel skuamosa yang melapisi saluran bronkus pada paru-paru. Proses perkembangan kanker ini biasanya diawali oleh perubahan pada sel epitel bronkus akibat paparan zat karsinogenik seperti asap rokok, polusi udara, atau bahan kimia berbahaya. Perubahan tersebut dapat

menyebabkan terjadinya metaplasia skuamosa, yang kemudian berkembang menjadi disampelasia dan pada akhirnya menjadi karsinoma invasif. Secara mikroskopis, LUSC ditandai dengan adanya keratinisasi, pembentukan keratin pearls, serta intercellular bridges yang merupakan ciri khas diferensiasi sel skuamosa (Lau et al., 2022). Karakteristik histomorfologis tersebut dapat diamati secara visual melalui pengamatan mikroskopis jaringan paru. Pada Gambar 2.1 (A) diperlihatkan adanya proses keratinisasi serta pembentukan keratin pearls atau mutiara keratin yang merupakan penanda utama karsinoma sel skuamosa. Selain itu, ciri khas berupa intercellular bridges (jembatan antar-sel) yang menandakan diferensiasi sel skuamosa juga tampak jelas pada Gambar 2.1 (B). Gambaran mikroskopis ini memberikan pemahaman mengenai kompleksitas struktur jaringan kanker LUSC yang akan dideteksi berdasarkan profil ekspresi gennya dalam penelitian ini.



Gambar 2. 1 Karakteristik Mikroskopis Kanker Paru Sel Skuamosa (LUSC).
(A) Keratinisasi (panah) dan keratin pearls (kepala panah). (B) Jembatan antar-sel atau intercellular bridges (panah).
(Sumber : Berezowska dkk., 2024)

Dari sisi epidemiologi, LUSC lebih sering ditemukan pada laki-laki dan individu dengan riwayat merokok dalam jangka waktu lama. Selain faktor merokok, beberapa faktor risiko lain yang berkontribusi terhadap perkembangan

kanker paru meliputi paparan polusi udara, bahan kimia industri, serta faktor genetik (Anggriani et al., 2022). Meskipun berbagai upaya pencegahan telah dilakukan, kanker paru masih sering terdiagnosis pada stadium lanjut karena gejala awal yang tidak spesifik. Kondisi ini menyebabkan tingkat kelangsungan hidup pasien relatif rendah apabila dibandingkan dengan jenis kanker lainnya.

Pada tingkat molekuler, LUSC memiliki karakteristik genomik yang kompleks. Berbagai penelitian menunjukkan bahwa LUSC sering dikaitkan dengan mutasi pada beberapa gen penting, seperti TP53, KEAP1, dan NFE2L2, yang berperan dalam regulasi siklus sel dan respons terhadap stres oksidatif. Selain itu, amplifikasi gen seperti FGFR1 dan perubahan pada jalur sinyal PI3K/AKT juga ditemukan pada sejumlah kasus LUSC (Luo et al., 2025). Perubahan genetik tersebut dapat menyebabkan deregulasi proses proliferasi sel, peningkatan resistensi terhadap apoptosis, serta perkembangan tumor yang lebih agresif.

Perkembangan teknologi genomik modern memungkinkan analisis kanker paru tidak hanya berdasarkan karakteristik histopatologis, tetapi juga melalui pendekatan molekuler seperti analisis ekspresi gen. Data ekspresi gen menggambarkan aktivitas biologis dari ribuan gen secara simultan sehingga dapat memberikan informasi yang lebih komprehensif mengenai kondisi sel (Wang et al., 2022). Pada jaringan kanker, pola ekspresi gen biasanya mengalami perubahan signifikan dibandingkan jaringan normal, baik berupa peningkatan (*up-regulation*) maupun penurunan ekspresi (*down-regulation*) pada sejumlah gen tertentu.

Melalui analisis bioinformatika dan pembelajaran mesin, pola ekspresi gen tersebut dapat dimanfaatkan untuk membedakan jaringan paru normal dan jaringan

kanker secara otomatis. Dataset genomik publik seperti *The Cancer Genome Atlas* (TCGA) menyediakan data ekspresi gen berskala besar yang dapat digunakan untuk mengembangkan model klasifikasi kanker berbasis kecerdasan buatan. Oleh karena itu, pemanfaatan data ekspresi gen pada penelitian ini diharapkan mampu membantu proses identifikasi kanker paru secara lebih cepat dan akurat melalui pendekatan komputasi. Dengan demikian, pemahaman mengenai karakteristik biologis dan molekuler LUSC menjadi landasan penting dalam penelitian ini, terutama dalam memanfaatkan pola ekspresi gen sebagai fitur utama dalam proses klasifikasi menggunakan metode *Multilayer Perceptron* (MLP).

2.3 Data Ekspresi Gen dan Normalisasi TPM

Ekspresi gen merupakan proses biologis di mana informasi genetik yang tersimpan dalam DNA ditranskripsikan menjadi RNA dan selanjutnya diterjemahkan menjadi protein. Tingkat ekspresi gen mencerminkan aktivitas biologis suatu sel karena protein yang dihasilkan berperan dalam berbagai proses seluler seperti proliferasi, diferensiasi, metabolisme, serta respons terhadap stres. Oleh karena itu, analisis ekspresi gen menjadi salah satu pendekatan penting dalam memahami karakteristik molekuler berbagai penyakit, termasuk kanker paru sel skuamosa (*Lung Squamous Cell Carcinoma* atau LUSC).

Pada sel kanker, pola ekspresi gen biasanya mengalami perubahan signifikan dibandingkan dengan sel normal. Perubahan tersebut dapat berupa peningkatan ekspresi (*up-regulation*) maupun penurunan ekspresi (*down-regulation*) pada sejumlah gen yang berperan dalam berbagai proses biologis yang berkaitan dengan perkembangan tumor. Perubahan pola ekspresi gen ini

menghasilkan karakteristik molekuler tertentu yang dikenal sebagai *gene expression signature*. Dengan memanfaatkan informasi tersebut, analisis ekspresi gen dapat digunakan untuk membedakan jaringan normal dan jaringan kanker secara lebih komprehensif dibandingkan hanya menggunakan data mutasi genetik (Liu et al., 2021; Wu et al., 2023).

Pengukuran ekspresi gen dalam penelitian genomik modern umumnya dilakukan menggunakan teknologi RNA *sequencing* (RNA-seq). RNA-seq merupakan metode yang memungkinkan pengukuran jumlah transkrip RNA secara langsung dengan tingkat sensitivitas yang tinggi. Teknologi ini mampu mengukur puluhan ribu gen secara simultan dalam satu sampel biologis sehingga menghasilkan representasi kuantitatif mengenai aktivitas gen dalam sel. Hasil pengukuran RNA-seq biasanya disajikan dalam bentuk matriks data yang terdiri dari ribuan fitur gen dan sejumlah sampel biologis, sehingga sangat sesuai untuk dianalisis menggunakan pendekatan bioinformatika dan pembelajaran mesin (Wang et al., 2022).

Dalam analisis data RNA-seq, salah satu tahap penting adalah proses normalisasi data. Nilai ekspresi gen mentah (*raw counts*) yang dihasilkan dari proses sekuensing tidak dapat dibandingkan secara langsung antar sampel karena adanya variasi teknis seperti perbedaan kedalaman sekuensing (*sequencing depth*) dan panjang gen (*gene length*). Jika tidak dikoreksi, variasi tersebut dapat menyebabkan bias dalam analisis sehingga perbedaan nilai ekspresi gen tidak sepenuhnya mencerminkan perbedaan biologis yang sebenarnya.

Salah satu metode normalisasi yang umum digunakan dalam analisis RNA-seq adalah *Transcripts Per Million* (TPM). Metode TPM bekerja dengan menyesuaikan nilai ekspresi gen berdasarkan panjang gen serta jumlah total transkrip dalam suatu sampel sehingga nilai ekspresi gen menjadi lebih proporsional dan dapat dibandingkan secara langsung antar sampel. Dibandingkan dengan metode normalisasi lain seperti RPKM atau FPKM, TPM menghasilkan nilai ekspresi yang lebih stabil dan lebih mudah diinterpretasikan dalam analisis komparatif antar sampel (Zhang et al., 2024; Huang et al., 2025).

Pada penelitian ini, data ekspresi gen yang digunakan telah tersedia dalam bentuk nilai yang telah dinormalisasi menggunakan metode TPM oleh *pipeline* bioinformatika pada basis data genomik yang digunakan. Oleh karena itu, penelitian ini tidak melakukan proses perhitungan TPM secara langsung, melainkan menggunakan data ekspresi gen yang telah dinormalisasi tersebut sebagai input dalam proses analisis selanjutnya.

Penggunaan data ekspresi gen yang telah dinormalisasi memungkinkan algoritma pembelajaran mesin untuk mempelajari pola hubungan antar gen secara lebih efektif. Dengan skala data yang lebih konsisten dan bias teknis yang telah diminimalkan melalui normalisasi TPM, model klasifikasi dapat lebih fokus pada variasi biologis yang relevan antara jaringan normal dan jaringan kanker. Hal ini menjadi dasar penting dalam pengembangan model klasifikasi berbasis *Multilayer Perceptron* (MLP) pada penelitian ini.

2.4 *Multilayer Perceptron*

Multilayer Perceptron (MLP) merupakan salah satu arsitektur *Artificial Neural Network* (ANN) yang bersifat *feed-forward*, yaitu informasi diproses secara berurutan dari lapisan input menuju satu atau lebih *hidden layer* dan diakhiri pada lapisan *output* tanpa adanya mekanisme umpan balik langsung. Setiap lapisan berperan melakukan transformasi terhadap data masukan sehingga jaringan mampu membentuk representasi fitur yang semakin abstrak. Secara teoritis, jaringan *feed-forward* memiliki kemampuan aproksimasi universal, sehingga mampu memodelkan hubungan non-linear yang kompleks pada data biomedis, termasuk data ekspresi gen yang umumnya mengandung interaksi antar gen yang tidak sederhana (Du et al., 2022).

Pada MLP, setiap neuron menerima keluaran dari lapisan sebelumnya, kemudian menghitung penjumlahan berbobot dan bias sebelum dilewatkan ke fungsi aktivasi. Proses ini dikenal sebagai forward *propagation*. Selanjutnya, pembelajaran dilakukan melalui *backpropagation*, yaitu mekanisme untuk menghitung gradien kesalahan terhadap seluruh parameter jaringan agar bobot dan bias dapat diperbarui secara iteratif hingga nilai *loss* semakin kecil. Dalam praktik modern, fungsi aktivasi ReLU banyak digunakan pada *hidden layer* karena mampu menghasilkan representasi yang jarang (*sparse*) dan mendukung proses pelatihan yang efisien, sedangkan pada tugas klasifikasi biner lapisan *output* umumnya menggunakan fungsi sigmoid untuk menghasilkan probabilitas kelas pada rentang 0 sampai 1 (Alkawaz et al., 2023; Shatravin et al., 2022).

Dalam analisis data ekspresi gen, penggunaan MLP dinilai relevan karena data transkriptomik umumnya berbentuk data tabular berdimensi sangat tinggi, tetapi tidak memiliki struktur spasial seperti citra maupun struktur urutan yang eksampelisit seperti teks. Oleh sebab itu, arsitektur yang paling umum digunakan untuk prediksi berbasis *gene expression* adalah MLP atau variasi jaringan *fully connected*. Tantangan utama pada data ini adalah jumlah fitur yang sangat besar, jumlah sampel yang relatif terbatas, serta tingginya risiko *overfitting*. Karena itu, proses normalisasi, seleksi fitur, regularisasi, dan penentuan *hyperparameter* menjadi faktor penting agar model mampu melakukan generalisasi dengan baik. (Gupta et al., 2022; Shen et al., 2022)

Beberapa penelitian menunjukkan bahwa MLP tetap kompetitif untuk tugas klasifikasi kanker berbasis ekspresi gen, meskipun performanya sangat dipengaruhi oleh ukuran dataset dan desain eksperimen. Hanczar et al. (2022) menunjukkan bahwa jaringan saraf berbasis ekspresi gen menjadi semakin kompetitif ketika ukuran data latih membesar, sementara pada kondisi sampel kecil strategi seperti *transfer learning* dapat membantu meningkatkan performa. Selain itu, Yang et al. (2022) mengembangkan MLP classifier berbasis 38-*gene signature* untuk subtipe kanker payudara dan melaporkan akurasi 93.56%, yang memperlihatkan bahwa kombinasi seleksi fitur biologis dan MLP dapat menghasilkan klasifikasi molekuler yang kuat. Dengan demikian, pada penelitian ini MLP sesuai digunakan karena input yang diproses berupa data ekspresi gen yang telah dinormalisasi dan diseleksi, sehingga jaringan dapat difokuskan pada gen-gen yang paling informatif untuk

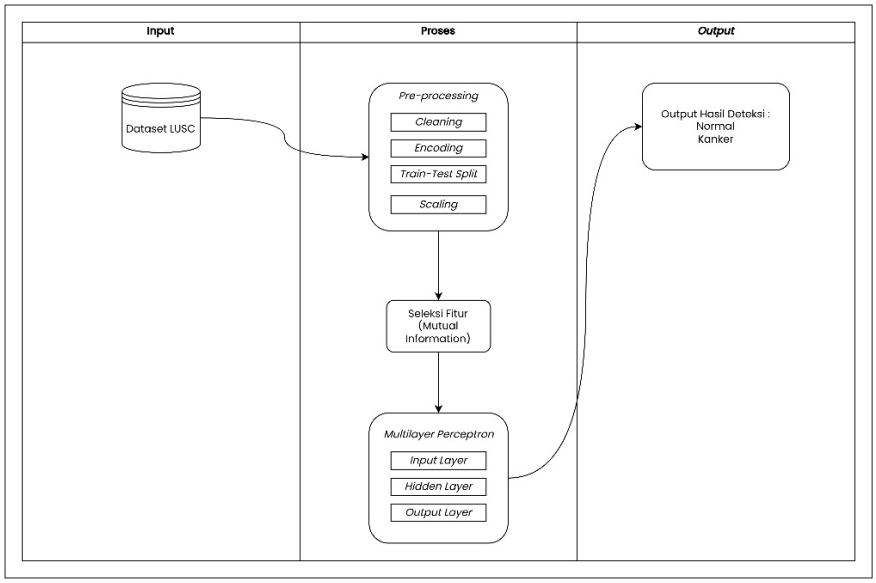
membedakan jaringan paru normal dan kanker LUSC. (Hanczar et al., 2022; Yang et al., 2022).

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Desain Sistem

Desain sistem pada penelitian ini menggambarkan alur proses klasifikasi data ekspresi gen mulai dari tahap input dataset hingga menghasilkan *output* berupa prediksi kelas. Perancangan sistem dilakukan secara terstruktur untuk memastikan setiap tahapan pengolahan data berjalan secara sistematis dan terintegrasi. Secara umum, sistem terdiri atas beberapa tahap utama, yaitu *preprocessing* data, seleksi fitur menggunakan Mutual Information, pelatihan model *Multilayer Perceptron* (MLP), serta proses klasifikasi untuk menghasilkan *output* prediksi. Setiap tahapan memiliki peran penting dalam meningkatkan kualitas data dan performa model, sehingga keseluruhan proses mampu menghasilkan sistem klasifikasi yang optimal. Alur desain sistem dalam penelitian ini ditunjukkan pada Gambar 3.1.



Gambar 3. 1 Desain penelitian

3.2 Pengumpulan Data

Bagian ini menguraikan proses pengumpulan dan penyiapan dataset ekspresi gen *Lung Squamous Cell Carcinoma* (LUSC) dari *The Cancer Genome Atlas* (TCGA) yang digunakan dalam penelitian. Dataset bertipe *RNA-sequencing* ini memiliki karakteristik *high-dimensional low-sample-size* (HDLSS) dan telah dinormalisasi menggunakan metode *Transcripts Per Million* (TPM). Mengingat distribusi jumlah sampel yang sangat tidak seimbang antara jaringan paru normal dan jaringan kanker, penelitian ini mengaplikasikan metode augmentasi data *Synthetic Minority Oversampling Technique* (SMOTE) secara khusus pada data pelatihan. Penerapan teknik ini bertujuan untuk menghasilkan sampel sintetis pada kelas minoritas melalui proses interpolasi, sehingga distribusi kelas menjadi lebih representatif dan mencegah model klasifikasi menjadi bias terhadap kelas mayoritas.

3.2.1 Eksperimen Data

Penelitian ini menggunakan dataset ekspresi gen *Lung Squamous Cell Carcinoma* (LUSC) yang bersumber dari portal genomik publik *The Cancer Genome Atlas* (TCGA) dan diakses melalui platform Kaggle (noepinefrin, 2024). Dataset TCGA-LUSC ini mencakup informasi profil ekspresi gen dari jaringan paru normal dan jaringan kanker paru subtipe LUSC. Seluruh data diperoleh menggunakan teknologi RNA *sequencing* (RNA-seq) dan telah melalui tahap pra-proses normalisasi dengan metode *Transcripts Per Million* (TPM). Hal ini memungkinkan nilai ekspresi gen untuk dibandingkan secara langsung dan konsisten antar seluruh sampel yang digunakan.

Secara keseluruhan dataset terdiri dari 551 sampel pasien, dengan rincian 49 sampel jaringan paru normal dan 502 sampel jaringan kanker LUSC. Setiap sampel memiliki 56.907 fitur gen yang merepresentasikan tingkat ekspresi dari masing-masing gen pada jaringan tersebut. Dengan demikian dataset memiliki struktur matriks berukuran 551 × 56.907, yang termasuk dalam kategori *high-dimensional low-sample-size* (HDLSS) karena jumlah fitur jauh lebih besar dibandingkan jumlah sampel. Karakteristik dataset yang digunakan dalam penelitian ini ditunjukkan pada Tabel 3.1.

Tabel 3. 1 Karakteristik Dataset TCGA-LUSC

Karakteristik	Nilai
Sumber dataset	TCGA (*The Cancer Genome Atlas*)
Total sampel	551
Jumlah Sampel normal	49
Jumlah Sampel kanker LUSC	502
Jumlah fitur gen	56.907
Tipe data	RNA-seq
Metode Normalisasi	TPM

Struktur dataset berbentuk matriks ekspresi gen di mana setiap baris merepresentasikan nama gen, sedangkan setiap kolom merepresentasikan sampel pasien. Nilai pada matriks tersebut menunjukkan tingkat ekspresi gen pada masing-masing sampel. Contoh sebagian struktur dataset ditunjukkan pada Tabel 3.2.

Tabel 3. 2 Contoh Struktur Dataset Ekspresi Gen

NO	Gene	Sample_1	Sample_2	Sample_3	Sample_4	Sample_5
1	Class	Normal	Normal	Normal	Tumor	Normal
2	A1BG	0,108403483	0,095187869	0,097042537	0,017848088	0,042861
3	A1BG-AS1	1.187325638	0,594016407	1.012277195	0,187729763	1.015939279
4	A1CF	0,001370356	0,032045614	0	0,003722766	0
5	A2M	459.7836081	648.0057316	444.8724975	57.23277815	751.7312734
6	A2M-AS1	0,486278098	1.080389631	1.047980517	1.369971025	0,780322401
7	ZYXP1	0	0	0	0	0
8	ZZEF1	0,512348	5.626855857	2.309259344	4.868548392	5.302770622

Pada dataset ini, label kelas ditentukan berdasarkan jenis jaringan paru yang diamati. Sampel dengan label normal menunjukkan jaringan paru sehat, sedangkan sampel dengan label *cancer* menunjukkan jaringan paru yang terdiagnosis sebagai kanker LUSC. Distribusi kelas pada dataset menunjukkan ketidakseimbangan jumlah sampel, di mana jumlah sampel kanker jauh lebih besar dibandingkan sampel normal. Kondisi ini dapat menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas. Oleh karena itu, pada penelitian ini diterapkan teknik *Synthetic Minority Oversampling Technique* (SMOTE) pada data pelatihan untuk menyeimbangkan distribusi kelas sebelum proses pelatihan model dilakukan.

3.2.2 Augmentasi Data

Dataset TCGA-LUSC yang digunakan dalam penelitian ini memiliki distribusi kelas yang tidak seimbang, di mana jumlah sampel kanker jauh lebih banyak dibandingkan sampel jaringan paru normal. Ketidakseimbangan kelas (*class imbalan*ce) dapat menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas sehingga performa deteksi pada kelas minoritas menjadi kurang optimal.

Untuk mengatasi permasalahan tersebut, penelitian ini menerapkan teknik *Synthetic MInority Oversampling Technique (SMOTE*) sebagai metode augmentasi data pada data pelatihan. SMOTE merupakan metode *oversampling* yang menghasilkan sampel sintetis pada kelas minoritas dengan memanfaatkan informasi dari tetangga terdekat (*K-Nearest Neighbors*) dalam ruang fitur.

Berbeda dengan metode *oversampling* sederhana yang hanya menduplikasi data yang ada, SMOTE menghasilkan sampel baru melalui proses interpolasi antara

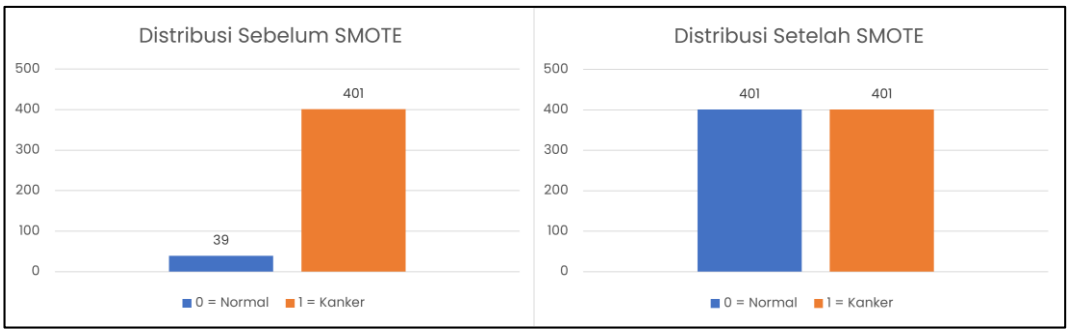
dua sampel dari kelas minoritas. Dengan demikian, distribusi data menjadi lebih seimbang tanpa menyebabkan *overfitting* akibat duplikasi data. Secara matematis, proses pembangkitan sampel sintetis pada SMOTE dapat dirumuskan sebagai berikut:

$$x_{new} = x_i + \lambda(x_{nn} - x_i) \tag{3.1}$$

Keterangan:
x_i: sampel dari kelas minoritas
x_{zi}: salah satu tetangga terdekat dari x_i
r: bilangan acak dalam rentang 0 sampai 1
x_{new}: sampel sintetis yang dihasilkan

Melalui mekanisme ini, sampel sintetis akan berada pada garis yang menghubungkan dua sampel minoritas sehingga distribusi data menjadi lebih representatif. Dalam penelitian ini, SMOTE diterapkan hanya pada data pelatihan untuk menjaga validitas proses evaluasi model dan mencegah terjadinya data leakage pada data pengujian.

Penerapan teknik augmentasi data ini diharapkan dapat meningkatkan kemampuan model dalam mempelajari pola ekspresi gen pada kelas minoritas sehingga performa klasifikasi antara jaringan paru normal dan kanker LUSC menjadi lebih seimbang.



Gambar 3. 2 Distribusi data sebelum dan setelah smote

Gambar 3.2 menunjukkan distribusi kelas sebelum dan sesudah dilakukan SMOTE. Distribusi sebelum resampling menunjukkan bahwa kelas 0 (Normal) berjumlah 39 data, sedangkan kelas 1 (Kanker) hanya berjumlah 401 data. Ketidakseimbangan ini dapat menyebabkan model cenderung bias terhadap kelas mayoritas. Sedangkan distribusi setelah dilakukan SMOTE menunjukkan bahwa kedua kelas memiliki jumlah data yang seimbang, masing masing 401 data untuk kelas 0 (Normal) dan 401 data untuk kelas 1 (Kanker). Untuk memperjelas alur komputasi dari setiap persamaan matematis yang digunakan dalam penelitian ini, akan diberikan sebuah ilustrasi perhitungan (*toy example*). Pada ilustrasi ini, diasumsikan terdapat sebuah data miniatur yang terdiri atas dua fitur gen (Gen A dan Gen B) serta dua sampel dari kelas minoritas (Jaringan Normal, y=0).

1. Sampel minoritas 1 (x_1): Gen A = 0,10 ; Gen B = 450,0
2. Sampel minoritas 2 (x_2): Gen A = 0,20 ; Gen B = 500,0

Berdasarkan data ilustrasi, kita akan membangkitkan satu sampel sintetis (x_{new}) dari kelas minoritas (Normal) menggunakan sampel $x_1$dan tetangga terdekatnya x_2. Jika ditentukan nilai bilangan acak $\lambda = 0{,}5$, maka proses augmentasi untuk masing-masing fitur adalah sebagai berikut:

Untuk Gen A:

$$x_{new(GenA)} = 0{,}10 + 0{,}5(0{,}20 - 0{,}10)$$

$$x_{new(GenA)} = 0{,}10 + 0{,}5(0{,}10)$$

$$x_{new(GenA)} = 0{,}10 + 0{,}05 = 0{,}15$$

Untuk Gen B:

$$x_{new(GenB)} = 450{,}0 + 0{,}5(500{,}0 - 450{,}0)$$

$$x_{new(GenB)} = 450{,}0 + 0{,}5(50{,}0)$$

$$x_{new(GenB)} = 450{,}0 + 25.0 = 475.0$$

Melalui interpolasi ini, diperoleh satu sampel sintetis baru kelas Normal dengan nilai fitur x (new) = [0,15 ; 475.0]. Sampel baru ini kemudian digabungkan ke dalam dataset pelatihan untuk menyeimbangkan distribusi kelas.

3.3 Preprocessing Data

Tahap *preprocessing* dilakukan untuk mempersiapkan dataset sebelum digunakan dalam proses pelatihan model klasifikasi. Pada tahap awal, dilakukan pemeriksaan terhadap kelengkapan dan konsistensi data, termasuk pengecekan nilai duplikat dan nilai hilang (*missing value*) yang berpotensi memengaruhi hasil analisis. Label kelas kemudian dikonversi ke dalam bentuk numerik untuk keperluan klasifikasi biner, di mana jaringan paru normal diberi label 0 dan jaringan kanker LUSC diberi label 1.

Meskipun data ekspresi gen telah dinormalisasi menggunakan metode *Transcripts Per Million* (TPM), proses standarisasi tetap dilakukan untuk menyamakan skala numerik antar fitur sebelum dimasukkan ke dalam model *Multilayer Perceptron* (MLP). Standarisasi dilakukan menggunakan metode *StandardScaler*, yang mentransformasikan setiap fitur sehingga memiliki rata-rata mendekati nol dan standar deviasi sebesar satu. Secara matematis, proses standarisasi dirumuskan sebagai:

$$z = \frac{x - \mu}{\sigma} \tag{3.2}$$

Keterangan
x : nilai asli suatu fitur
μ : rata-rata (mean) fitur pada data pelatihan
σ : standar deviasi fitur tersebut

Standarisasi dilakukan untuk menyamakan skala antar fitur sehingga setiap gen memiliki rata-rata mendekati nol dan standar deviasi sebesar satu sebelum dimasukkan ke dalam model MLP. Selanjutnya, dataset dibagi menjadi data pelatihan dan data pengujian menggunakan metode stratified train-test sampelit untuk mempertahankan proporsi distribusi kelas pada kedua subset data. Pembagian ini memastikan bahwa evaluasi model dilakukan pada data yang tidak terlibat dalam proses pelatihan. Adapun contoh perubahan nilai data akibat proses standarisasi ini diilustrasikan pada Gambar 3.3, di mana rentang nilai antar fitur gen menjadi lebih seragam dan terpusat.

Data Sebelum Scaling

	A1BG	A1BG-AS1	A1CF	A2M	A2M-AS1
92	0.066671	0.486099	0.004635	40.699838	1.218447
216	0.185907	1.455131	0.004432	9.472636	0.223268
440	0.005298	0.129194	0.000000	24.710187	0.397015
420	0.028023	0.401933	0.000000	20.641789	0.102427
105	0.183823	1.096260	0.004036	78.140631	0.061885

Data Setelah Scaling

	A1BG	A1BG-AS1	A1CF	A2M	A2M-AS1
92	-0.433616	-0.552014	-0.070801	-0.451602	1.410681
216	0.971961	0.962564	-0.071573	-0.583286	-1.094260
440	-1.157081	-1.109850	-0.088365	-0.519030	-0.656927
420	-0.889202	-0.683563	-0.088365	-0.536186	-1.398426
105	0.947391	0.401656	-0.073072	-0.293715	-1.500474

Gambar 3. 3 Perbandingan cuplikan data ekspresi gen sebelum dan sesudah standarisasi

Untuk keperluan ilustrasi ini, diasumsikan berdasarkan perhitungan pada seluruh data pelatihan, fitur Gen A memiliki nilai rata-rata $(\mu) = 0{,}50$ dan standar deviasi $(\sigma) = 0{,}25$. Sementara itu, fitur Gen B memiliki nilai rata-rata $(\mu) = 400{,}0$ dan standar deviasi $(\sigma) = 100{,}0$, Dalam proses ini akan menstandarisasi nilai dari sampel sintetis $x_{new} = [0{,}15; 475{,}0]$ yang diperoleh dari tahap

augmentasi sebelumnya. Berdasarkan Persamaan 3.2, perhitungannya adalah sebagai berikut:

Untuk fitur Gen A:

$$z_{GenA} = \frac{0{,}15 - 0{,}50}{0{,}25}$$

$$z_{GenA} = \frac{-0{,}35}{0{,}25} = -1.40$$

Untuk fitur Gen B:

$$z_{GenB} = \frac{475.0 - 400{,}0}{100{,}0}$$

$$z_{GenB} = \frac{75.0}{100{,}0} = 0{,}75$$

Melalui proses standarisasi ini, sampel tersebut kini memiliki vektor fitur dengan skala yang baru, yaitu:

$$Z = [-1.40;\ 0{,}75]$$

Vektor nilai yang telah distandarisasi inilah yang selanjutnya akan diproses dalam tahapan seleksi fitur dan menjadi masukan (input) bagi model *Multilayer Perceptron* (MLP).

3.4 Seleksi Fitur

Dataset ekspresi gen yang digunakan dalam penelitian ini memiliki jumlah fitur yang sangat besar (56.907 gen) dibandingkan jumlah sampel (551 sampel). Kondisi ini termasuk dalam kategori high-dimensional, *low-sample-size* (HDLSS) yang berpotensi menyebabkan *overfitting* serta menurunkan kemampuan generalisasi model. Oleh karena itu, diperlukan teknik seleksi fitur untuk

mengurangi dimensi data dengan tetap mempertahankan informasi yang relevan terhadap proses klasifikasi.

Pada penelitian ini, metode *Mutual Information* (MI) digunakan untuk melakukan seleksi fitur. Pemilihan metode ini didasarkan pada kemampuannya dalam mengukur tingkat ketergantungan antara suatu fitur dan label kelas tanpa mengasumsikan bentuk hubungan linear tertentu. Berbeda dengan metode statistik seperti ANOVA atau *F-test* yang berfokus pada perbedaan rata-rata, MI mampu menangkap hubungan non-linear antara gen dan kelas target, sehingga lebih sesuai untuk data ekspresi gen yang memiliki pola biologis kompleks.

Secara matematis, *Mutual Information* antara dua variabel acak X (fitur gen) dan Y (label kelas) dirumuskan sebagai:

$$I(X;Y) = \sum_{y \in Y} \int_X p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) dx \qquad (3.3)$$

Keterangan
$I(X;Y)$: nilai *Mutual Information* antara fitur dan kelas,
X: Variabel acak kontinu yang merepresentasikan nilai ekspresi suatu gen.
Y: Variabel acak diskret yang merepresentasikan label kelas (0 = normal, 1 = kanker LUSC).
$p(x,y)$: Fungsi kepadatan probabilitas gabungan antara ekspresi gen dan kelas.
$p(x)$: Fungsi kepadatan probabilitas marginal dari ekspresi gen.
$p(y)$: Probabilitas marginal dari kelas target.

Karena data ekspresi gen bersifat kontinu, perhitungan MI pada sistem komputasi umumnya melibatkan estimasi fungsi kepadatan probabilitas (integral). Namun, untuk ilustrasi matematis ini, nilai fitur Gen A disederhanakan ke dalam bentuk diskret: “Tinggi (T)” dan “Rendah (R)”. Diasumsikan terdapat 4 sampel:

1. 2 sampel Normal (y = 0) → keduanya Rendah (R)
2. 2 sampel Kanker (y = 1) → 1 Rendah (R), 1 Tinggi (T)

Sehingga:

$$P(R) = \frac{3}{4}, P(T) = \frac{1}{4}$$

a. Entropi Gen A $H(X)$

$$H(X) = -(P(R)\log_2 P(R) + P(T)\log_2 P(T))$$

$$H(X) = -\left(\frac{3}{4}\log_2 \frac{3}{4} + \frac{1}{4}\log_2 \frac{1}{4}\right)$$

$$H(X) \approx -(0{,}75(-0{,}415) + 0{,}25(-2))$$

$$H(X) \approx 0{,}311 + 0{,}5 = 0{,}811 \text{ bit}$$

b. Entropi Kondisional $H(X|Y)$

Untuk kelas Normal (y = 0):

$$H(X \mid y = 0) = -(1\log_2 1) = 0 \text{ bit}$$

Untuk kelas Kanker (y = 1):

$$H(X \mid y = 1) = -(0{,}5\log_2 0{,}5 + 0{,}5\log_2 0{,}5) = 1 \text{ bit}$$

Total entropi kondisional:

$$H(X \mid Y) = 0{,}5(0) + 0{,}5(1) = 0{,}5 \text{ bit}$$

c. *Mutual Information* $(I(X;Y))$

$$I(X;Y) = H(X) - H(X \mid Y)$$

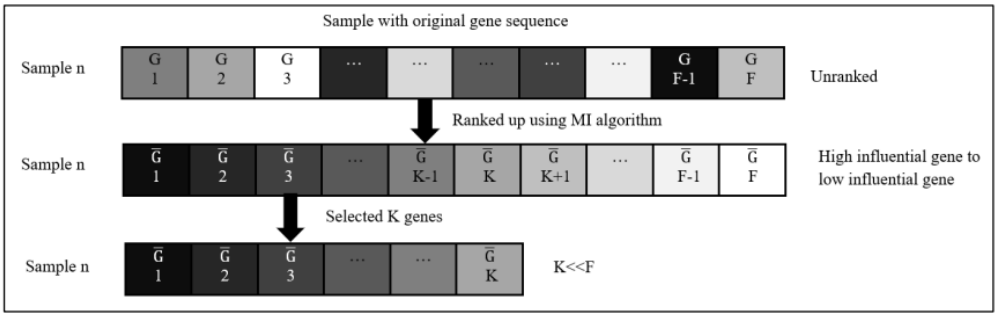
$$I(X;Y) = 0{,}811 - 0{,}5 = 0{,}311 \text{ bit}$$

Proses ini diulang untuk seluruh gen. Pada ilustrasi ini, diasumsikan Gen A dan Gen B memiliki nilai MI tertinggi (terpilih sebagai Top-2 *Feature* s).Dengan demikian, nilai *Mutual Information* antara Gen A dan kelas target adalah 0,311 bit.

Mutual Information digunakan untuk mengukur seberapa besar ketergantungan antara suatu fitur gen dengan label kelas. Semakin tinggi nilainya, semakin relevan fitur tersebut dalam proses klasifikasi. Nilai MI bersifat non-

negatif, dan semakin besar nilainya menunjukkan bahwa fitur tersebut memiliki kontribusi yang lebih signifikan terhadap proses klasifikasi. Proses seleksi fitur dilakukan pada data pelatihan untuk menghindari *data leakage*, dengan langkah-langkah sebagai berikut:

1. Menghitung nilai MI setiap gen terhadap label kelas.
2. Mengurutkan gen berdasarkan nilai MI secara menurun.
3. Memilih sejumlah K gen dengan nilai MI tertinggi sebagai *top-K feature s*.



Gambar 3. 4 Ilustrasi Proses Seleksi Fitur Menggunakan Mutual Information
(Sumber : Tabassum dkk., 2024)

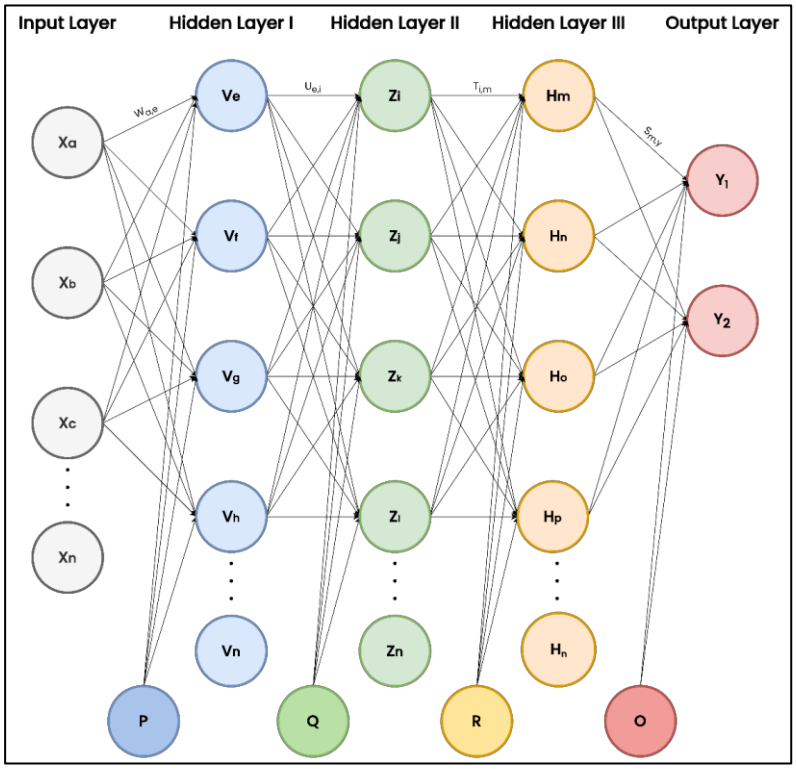
Ilustrasi proses seleksi fitur berbasis *Mutual Information* ditunjukkan pada Gambar 3.3. Seluruh gen awalnya berada dalam kondisi tidak terurut (*unranked*). Setelah dihitung nilai MI terhadap label kelas, gen-gen tersebut diurutkan dari yang paling berpengaruh hingga yang paling rendah kontribusinya. Selanjutnya dipilih sejumlah K gen dengan nilai MI tertinggi (K << F), yang kemudian digunakan sebagai input model.

Pendekatan ini juga digunakan dalam penelitian Tabassum et al. (2024), di mana jumlah fitur terbaik ditentukan melalui pendekatan *trial-and-error* untuk memperoleh performa klasifikasi optimal. Dalam penelitian ini, variasi jumlah fitur yang diuji meliputi seluruh fitur (tanpa seleksi), serta beberapa skenario top-K fitur

untuk menganalisis pengaruh jumlah fitur terhadap performa model. Reduksi jumlah fitur ini diharapkan mampu menurunkan kompleksitas komputasi, mempercepat proses pelatihan, serta mengurangi risiko *overfitting* pada model MLP.

3.5 Arsitektur *Multi-Layer Perceptron* (MLP)

Setelah proses seleksi fitur menggunakan metode *Mutual Information* (MI) dilakukan dan diperoleh sejumlah fitur gen terpilih (top-K *feature* s), tahap selanjutnya adalah melakukan proses klasifikasi menggunakan model *Multilayer Perceptron* (MLP). MLP merupakan salah satu jenis jaringan saraf tiruan (artificial neural network) yang bersifat *feedforward* dan mampu memodelkan hubungan non-linear antara fitur masukan dan label kelas. Model ini dipilih karena kemampuannya dalam menangani data berdimensi tinggi serta efektif dalam permasalahan klasifikasi biner, seperti pada penelitian ini yang bertujuan mengklasifikasikan jaringan paru menjadi kategori normal atau kanker LUSC.



Gambar 3. 5 Arsitektur *Multilayer Perceptron*

Untuk mempermudah pemahaman struktur jaringan yang digunakan, arsitektur MLP pada penelitian ini dapat divisualisasikan dalam bentuk diagram seperti pada Gambar 3.5, yang menunjukkan hubungan antara lapisan input, tiga lapisan tersembunyi, dan lapisan *output* dalam proses klasifikasi data ekspresi gen.

1. Tahapan *Feedforward Propagation*

Proses *feedforward propagation* merupakan tahapan awal dalam pelatihan jaringan *Multilayer Perceptron* (MLP), di mana data masukan diproses secara berurutan dari lapisan input menuju lapisan *output* untuk menghasilkan nilai prediksi. Pada tahap ini, setiap neuron menerima sinyal dari neuron pada lapisan sebelumnya, kemudian melakukan transformasi linier melalui operasi perkalian bobot dan penambahan bias. Hasil transformasi tersebut selanjutnya diproses menggunakan fungsi aktivasi untuk menghasilkan *output* pada setiap lapisan.

Misalkan vektor input hasil seleksi fitur dinyatakan sebagai:

$$X = [x_1, x_2, x_3, \dots, x_K]^T \tag{3.4}$$

dengan K merupakan jumlah fitur gen terpilih.

Setelah proses seleksi fitur dilakukan, sejumlah gen terbaik akan disusun menjadi vektor input yang selanjutnya digunakan sebagai masukan pada model MLP. Proses *feedforward propagation* pada arsitektur MLP dimulai dari aliran data pada lapisan input menuju *hidden layer* pertama, di mana vektor fitur hasil seleksi *Mutual Information* diproses melalui transformasi linier menggunakan bobot dan bias kemudian dilewatkan ke fungsi aktivasi ReLU, sebagaimana pada persamaan 3.5:

$$V_e = ReLU\left(\sum_{a=1}^{A} X_a \cdot W_{a,e} + P_e\right) \tag{3.5}$$

Keterangan:
X_a= Nilai input neuron ke-a
$W_{a,e}$= Bobot antara neuron input $X_a$dan neuron *hidden layer* 1 V_e
P_e= Bias pada neuron V_e
A= Jumlah total neuron pada lapisan input
V_e= *Output* neuron ke-e pada *hidden layer* 1

Pada tahap ini, setiap nilai input X_a dikalikan dengan bobot yang terhubung ke masing-masing neuron pada *hidden layer* pertama, kemudian seluruh hasil perkalian tersebut dijumlahkan dan ditambahkan nilai bias. Hasil penjumlahan berbobot tersebut selanjutnya diproses menggunakan fungsi aktivasi ReLU untuk menghasilkan keluaran pada lapisan tersembunyi pertama. Lapisan ini berperan dalam menangkap pola awal dari kombinasi gen-gen yang menjadi masukan model.

Untuk keperluan ilustrasi, diasumsikan *Hidden Layer* 1 terdiri dari 2 neuron (V_1dan V_2). Parameter bobot $(W_{a,e})$yang menghubungkan 2 input ke 2 neuron tersebut, beserta bias (P_e), diinisialisasi secara acak dengan nilai berikut:

a. Bobot menuju Neuron 1: $W_{1,1} = -0{,}5$ dan $W_{2,1} = 0{,}8$. Bias $P_1 = 0{,}1$.

b. Bobot menuju Neuron 2: $W_{1,2} = 0{,}2$ dan $W_{2,2} = -0{,}4$. Bias $P_2 = 0{,}2$.

Berdasarkan Persamaan 3.5, proses penjumlahan berbobot dan aktivasi menggunakan fungsi ReLU dihitung sebagai berikut:

a. Pada Neuron 1 (V_1):

$$z_1 = (X_1 \cdot W_{1,1}) + (X_2 \cdot W_{2,1}) + P_1$$

$$z_1 = (-1.40 \cdot -0{,}5) + (0{,}75 \cdot 0{,}8) + 0{,}1$$

$$z_1 = 0{,}70 + 0{,}60 + 0{,}1 = 1.40$$

$$V_1 = \text{ReLU}(1{,}40) = \max(0{,}1{,}40) = 1.40$$

b. Pada Neuron 2 (V_2):

$$V_2 = \text{ReLU}(-0{,}38) = \max(0, -0{,}38) = 0$$

Keluaran dari *Hidden Layer* 1 untuk sampel ini adalah vektor:

$$V = [1.40; 0]$$

Vektor keluaran V ini selanjutnya akan menjadi masukan bagi *Hidden Layer* berikutnya.Pada *hidden layer* kedua, keluaran dari *hidden layer* pertama dijadikan masukan untuk neuron-neuron pada lapisan berikutnya. Setiap nilai *output* dikalikan dengan bobot, dijumlahkan, lalu ditambahkan bias dan diproses menggunakan fungsi aktivasi ReLU. *Output* dari *hidden layer* pertama, yaitu V_e, selanjutnya menjadi masukan bagi *hidden layer* kedua. Pada tahap ini, setiap *output* neuron pada *hidden layer* pertama dikalikan dengan bobot yang terhubung ke

masing-masing neuron pada *hidden layer* kedua, kemudian seluruh hasil perkalian tersebut dijumlahkan dan ditambahkan nilai bias. Hasil penjumlahan berbobot tersebut selanjutnya diproses menggunakan fungsi aktivasi ReLU untuk menghasilkan keluaran pada *hidden layer* kedua, sebagaimana persamaan 3.6:

$$Z_i = ReLU\left(\sum_{e=1}^{E} V_e \cdot U_{e,i} + Q_i\right) \tag{3.6}$$

eterangan:
V_e= *Output* neuron ke-e pada *hidden layer* 1
$U_{e,i}$= Bobot antara neuron $V_e$dan neuron *hidden layer* 2 Z_i
Q_i= Bias pada neuron Z_i
E= Jumlah neuron pada *hidden layer* 1
Z_i= *Output* neuron ke-i pada *hidden layer* 2

Keluaran dari lapisan tersembunyi pertama, yaitu vektor $V = [1{,}40; 0]$, diteruskan sebagai masukan ke *Hidden Layer* 2. Asumsikan parameter bobot $(U_{e,i})$dan bias (Q_i)pada lapisan ini bernilai:

a. Bobot menuju Neuron 1 (Z_1): $U_{1,1} = 0{,}3$dan $U_{2,1} = -0{,}2$. Bias $Q_1 = -0{,}1$.

b. Bobot menuju Neuron 2 (Z_2): $U_{1,2} = -0{,}1$dan $U_{2,2} = 0{,}5$. Bias $Q_2 = 0{,}3$.

Berdasarkan Persamaan 3.6, perhitungannya adalah:

1) Pada Neuron 1 (Z_1):

$$input_{(Z1)} = (V_1 \cdot U_{1,1}) + (V_2 \cdot U_{2,1}) + Q_1$$

$$input_{(Z1)} = (1.40 \cdot 0{,}3) + (0 \cdot -0{,}2) - 0{,}1$$

$$input_{(Z1)} = 0{,}42 + 0 - 0{,}1 = 0{,}32$$

$$Z_1 = \text{ReLU}(0{,}32) = 0{,}32$$

2) Pada Neuron 2 (Z_2):

$$Z_2 = \text{ReLU}(0{,}16) = 0{,}16$$

Keluaran dari *Hidden Layer* 2 untuk sampel ini adalah vector $Z = [0{,}32; 0{,}16]$.

Pada *hidden layer* ketiga, keluaran dari *hidden layer* kedua diproses kembali dengan mekanisme yang sama, yaitu penjumlahan berbobot dan fungsi aktivasi ReLU. Pada tahap ini, keluaran dari *hidden layer* kedua, yaitu Z_i, diteruskan sebagai masukan ke *hidden layer* ketiga. Setiap *output* neuron pada *hidden layer* kedua dikalikan dengan bobot yang terhubung ke neuron pada *hidden layer* ketiga, kemudian seluruh hasil perkalian tersebut dijumlahkan dan ditambahkan nilai bias. Hasil penjumlahan berbobot ini selanjutnya diproses menggunakan fungsi aktivasi ReLU untuk menghasilkan keluaran pada *hidden layer* ketiga, sebagaimana persamaan 3.7:

$$H_m = ReLU\left(\sum_{i=1}^{I} Z_i \cdot T_{i,m} + R_j\right) \tag{3.7}$$

Keterangan:
Z_i= *Output* neuron ke-i pada *hidden layer* 2
$T_{i,m}$= Bobot antara neuron $Z_i$dan neuron *hidden layer* 3 H_j
R_j= Bias pada neuron H_j
I= Jumlah neuron pada *hidden layer* 2
H_m= *Output* neuron ke-j pada *hidden layer* 3

Dengan mekanisme perhitungan matriks yang identik dengan lapisan-lapisan sebelumnya, vektor keluaran $Z = [0{,}32; 0{,}16]$dikalikan dengan matriks bobot $(T_{i,m})$dan dijumlahkan dengan bias (R_j)pada *Hidden Layer* 3. Diasumsikan hasil akhir setelah melewati fungsi aktivasi ReLU pada *Hidden Layer* 3 menghasilkan vektor keluaran berikut:

$$H = \begin{bmatrix} 0{,}60 \\ 0{,}40 \end{bmatrix}$$

Vektor H inilah yang merupakan representasi fitur paling abstrak (terdalam) dari jaringan dan siap dilewatkan menuju lapisan *output* untuk proses pengambilan keputusan klasifikasi.

Keluaran dari *hidden layer* ketiga, yaitu H_m, selanjutnya diteruskan ke lapisan *output*. Pada tahap ini, setiap *output* neuron pada *hidden layer* ketiga dikalikan dengan bobot yang terhubung ke neuron *output*, kemudian seluruh hasil perkalian tersebut dijumlahkan dan ditambahkan nilai bias. Hasil penjumlahan berbobot tersebut selanjutnya diproses menggunakan fungsi aktivasi sigmoid $\sigma(\cdot)$ untuk menghasilkan nilai probabilitas $\hat{y}$ dalam rentang [0, 1], sebagaimana persamaan 3.8:

$$\hat{y} = \sigma\left(\sum_{j=1}^{J} H_m \cdot S_m + b\right) \tag{3.8}$$

Keterangan:

H_m= *Output* neuron ke-j pada *hidden layer* 3
S_m= Bobot antara neuron $H_j$dan *output*
b= Bias pada neuron *output*
J= Jumlah neuron pada *hidden layer* 3
σ= Fungsi aktivasi sigmoid
$\hat{y}$= Probabilitas prediksi kelas

Nilai $\hat{y}$ merepresentasikan probabilitas suatu sampel termasuk ke dalam kelas kanker LUSC (label 1). Semakin mendekati nilai 1, maka semakin besar kemungkinan sampel tersebut diklasifikasikan sebagai kanker, sedangkan nilai yang mendekati 0 menunjukkan bahwa sampel termasuk dalam kelas normal. Vektor keluaran dari lapisan sebelumnya, yakni $H = [0{,}60; 0{,}40]$, kini dimasukkan ke lapisan *output*. Pada lapisan ini hanya terdapat satu neuron. Diasumsikan bobot yang menghubungkan *Hidden Layer* 3 ke neuron *output* adalah $S_1 = 0{,}5$dan $S_2 =$

$-0{,}3$, dengan nilai bias $b = 0{,}1$. Berdasarkan Persamaan 3.8, penjumlahan berbobot dan aktivasi sigmoidnya dihitung sebagai berikut:

a. Perhitungan Nilai z_{out}

$$z_{out} = (H_1 \cdot S_1) + (H_2 \cdot S_2) + b$$

$$z_{out} = (0{,}60 \cdot 0{,}5) + (0{,}40 \cdot -0{,}3) + 0{,}1$$

$$z_{out} = 0{,}30 - 0{,}12 + 0{,}1 = 0{,}28$$

b. Fungsi Aktivasi Sigmoid

$$\hat{y} = \frac{1}{1 + e^{-z_{out}}}$$

$$\hat{y} = \frac{1}{1 + e^{-0{,}28}} \approx \frac{1}{1 + 0{,}755} \approx 0{,}57$$

Pada lapisan *output*, keluaran dari *hidden layer* ketiga dikalikan dengan bobot menuju neuron *output*, kemudian dijumlahkan dengan bias. Hasilnya diproses menggunakan fungsi sigmoid untuk memperoleh probabilitas prediksi kelas. Sebagai keputusan klasifikasi biner, nilai probabilitas $\hat{y}$ dapat dikonversi menjadi label kelas menggunakan ambang batas (threshold) 0,5, sebagaimana persamaan 3.9:

$$\hat{Y} = \begin{cases} 1, & \text{jika } \hat{y} \geq 0{,}5 \\ 0, & \text{jika } \hat{y} < 0{,}5 \end{cases} \tag{3.9}$$

di mana $\hat{Y} = 1$ menunjukkan kelas kanker LUSC, dan $\hat{Y} = 0$ menunjukkan kelas normal.

$$Y = 1, \text{karena } 0{,}57 \geq 0{,}5$$

Nilai probabilitas prediksi $(\hat{y})$yang dihasilkan adalah 0,57. Berdasarkan hasil komputasi *feedforward* ini, model mengklasifikasikan sampel tersebut sebagai kelas Kanker LUSC (Label 1).

Setelah diperoleh nilai probabilitas prediksi, langkah berikutnya adalah mengubah probabilitas tersebut menjadi label kelas menggunakan ambang batas 0,5. Jika probabilitas lebih besar atau sama dengan 0,5, maka sampel diklasifikasikan sebagai kelas kanker, sedangkan jika kurang dari 0,5 maka diklasifikasikan sebagai kelas normal.

Setelah proses *feedforward propagation* menghasilkan nilai prediksi berupa probabilitas $\hat{y}$, tahap selanjutnya adalah menghitung nilai kesalahan (*loss*) antara hasil prediksi model dan label sebenarnya. Pada penelitian ini, karena permasalahan yang dihadapi merupakan klasifikasi biner (normal dan kanker LUSC) dengan fungsi aktivasi sigmoid pada lapisan *output*, maka digunakan fungsi *Binary Cross Entropy* (BCE) sebagai fungsi *loss*. Fungsi *Binary Cross Entropy* mengukur perbedaan antara probabilitas prediksi $\hat{y}$ dan label aktual y, di mana nilai *loss* akan semakin kecil apabila prediksi mendekati nilai sebenarnya, sebagaimana persamaan 3.10:

$$L(y, \hat{y}) = -[y\log(\hat{y}) + (1-y)\log(1-\hat{y})] \tag{3.10}$$

Untuk keseluruhan data dengan jumlah sampel sebanyak n, fungsi *loss* dirumuskan sebagai rata-rata dari seluruh sampel, sebagaimana persamaan 3.11:

$$L = -\frac{1}{n}\sum_{i=1}^{n}[y_i\log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \tag{3.11}$$

Keterangan:
L: nilai *loss* rata-rata
n: jumlah sampel dalam satu batch atau dataset
y_i: label aktual sampel ke-i (0 untuk normal, 1 untuk kanker)
$\hat{y}_i$: probabilitas prediksi model pada sampel ke-i
log: fungsi logaritma natural

Fungsi *Binary Cross Entropy* dipilih karena sesuai dengan karakteristik *output* sigmoid yang menghasilkan nilai probabilitas dalam rentang 0 hingga 1. Selain itu, BCE memiliki sifat konveks pada klasifikasi logistik sehingga memudahkan proses optimisasi berbasis gradien. *Backpropagation*

Pada ilustrasi di atas, model memprediksi probabilitas $\hat{y} = 0{,}57$ (Kanker LUSC). Namun, label aktual dari sampel tersebut sebenarnya adalah jaringan Normal ($y = 0$). Terdapat kesalahan prediksi yang dilakukan oleh model.Untuk mengukur seberapa besar kesalahan tersebut, digunakan fungsi *Binary Cross-Entropy* (BCE) berdasarkan Persamaan 3.10, Perhitungannya menggunakan logaritma natural:

$$L(\hat{y}, y) = -y\ln(\hat{y}) - (1-y)\ln(1-\hat{y})$$
$$L = -0 \cdot \ln(0{,}57) - (1-0) \cdot \ln(1-0{,}57)$$
$$L = 0 - 1 \cdot \ln(0{,}43)$$
$$L = -(-0{,}844) = 0{,}844$$

Nilai *loss* sebesar 0,844 menunjukkan besarnya penalti akibat kesalahan klasifikasi pada sampel tersebut. Semakin besar tingkat keyakinan model terhadap prediksi yang salah (misalnya memprediksi 0,99 untuk sampel normal), maka nilai *loss* akan semakin melonjak tinggi secara eksponensial.

Setelah nilai *loss* dihitung menggunakan fungsi *Binary Cross Entropy*, tahap berikutnya adalah melakukan proses *backpropagation*. Proses ini bertujuan untuk menghitung gradien fungsi *loss* terhadap seluruh parameter jaringan, yaitu bobot dan bias, dengan menerapkan aturan turunan berantai (*chain rule*) dari lapisan *output* hingga ke lapisan-lapisan tersembunyi sebelumnya. Gradien yang diperoleh

menunjukkan arah perubahan parameter yang dapat menurunkan nilai *loss* pada iterasi berikutnya. Pada lapisan *output*, karena digunakan fungsi aktivasi sigmoid yang dikombinasikan dengan *Binary Cross Entropy*, sebagaimana persamaan 3.12:

$$\delta^{(4)} = \hat{y} - y \tag{3.12}$$

Keterangan:
$\delta^{(4)}$: nilai error pada lapisan *output*
$\hat{y}$: probabilitas prediksi model
y: label aktual sampel

Berdasarkan nilai error tersebut, gradien terhadap bobot dan bias pada lapisan *output* dihitung sebagaimana persamaan 3.13 dan 3.14:

$$\frac{\partial L}{\partial W^{(4)}} = \delta^{(4)} A^{(3)} \tag{3.13}$$

$$\frac{\partial L}{\partial b^{(4)}} = \delta^{(4)} \tag{3.14}$$

Keterangan:
$\frac{\partial L}{\partial W^{(4)}}$: gradien *loss* terhadap bobot pada *output layer*
$\frac{\partial L}{\partial b^{(4)}}$: gradien *loss* terhadap bias pada *output layer*
$A^{(3)}$: *output* dari *hidden layer* ketiga

Berdasarkan hasil komputasi *feedforward*, diperoleh nilai probabilitas prediksi $\hat{y} = 0{,}57$ sedangkan label aktual adalah $y = 0$, Sesuai Persamaan 3.12, nilai error pada lapisan *output* (δ_{out})dihitung sebagai berikut:

$$\delta_{out} = \hat{y} - y$$

$$\delta_{out} = 0{,}57 - 0 = 0{,}57$$

Nilai δ_{out}positif ini mengindikasikan bahwa tebakan model terlalu tinggi dan harus diturunkan. Selanjutnya, gradien (tingkat perubahan) terhadap bobot

(S_m)dan bias (b)pada lapisan *output* dihitung menggunakan Persamaan 3.13 dan 3.14. Mengingat input dari lapisan sebelumnya adalah $H = [0{,}60; 0{,}40]$:

a. Gradien terhadap bobot S_1

$$\frac{\partial L}{\partial S_1} = \delta_{out} \cdot H_1$$

$$\frac{\partial L}{\partial S_1} = 0{,}57 \cdot 0{,}60 = 0{,}342$$

b. Gradien terhadap bobot S_2

$$\frac{\partial L}{\partial S_2} = \delta_{out} \cdot H_2$$

$$\frac{\partial L}{\partial S_2} = 0{,}57 \cdot 0{,}40 = 0{,}228$$

c. Gradien terhadap bias b

$$\frac{\partial L}{\partial b} = \delta_{out} = 0{,}57$$

Nilai-nilai gradien ini menunjukkan seberapa besar kontribusi masing-masing parameter terhadap kesalahan total, dan arah mana yang harus diambil untuk memperbaikinya. Setelah nilai error pada lapisan *output* diperoleh, langkah berikutnya adalah menghitung gradien *loss* terhadap bobot pada *output layer*. Gradien ini menunjukkan seberapa besar kontribusi masing-masing bobot terhadap error yang dihasilkan. Selanjutnya, perhitungan error dilanjutkan ke lapisan tersembunyi dengan menerapkan aturan rantai (chain rule), yaitu mengalikan error dari lapisan berikutnya dengan turunan fungsi aktivasi pada lapisan tersebut, sebagaimana persamaan 3.15:

$$\delta^{(l)} = \left(W^{(l+1)T}\delta^{(l+1)}\right) \odot f'(Z^{(l)}) \tag{3.15}$$

Keterangan:
$\delta^{(l)}$: error pada *hidden layer* ke-l
$W^{(l+1)T}$: transpose matriks bobot dari *layer* berikutnya
$\delta^{(l+1)}$: error dari *layer* berikutnya
$f'(Z^{(l)})$: turunan fungsi aktivasi
$\odot$: perkalian elemen-per-elemen (*element-wise Multiplication*)

Setelah error pada *output layer* diperoleh, error tersebut disebarkan kembali ke lapisan tersembunyi menggunakan aturan rantai. Pada contoh ini, perhitungan dilakukan pada *hidden layer* ketiga, karena lapisan ini berada tepat sebelum *output layer* sehingga paling mudah ditelusuri. Karena pada penelitian ini *hidden layer* menggunakan fungsi aktivasi ReLU, sebagaimana persamaan 3.16:

$$f'(x) = \begin{cases} 1, & \text{jika } x > 0 \\ 0, & \text{jika } x \leq 0 \end{cases} \tag{3.16}$$

Keterangan:
$f'(x)$: turunan fungsi aktivasi ReLU
x : nilai sebelum aktivasi

Pada *hidden layer*, fungsi aktivasi yang digunakan adalah ReLU. Oleh karena itu, untuk proses *backpropagation* perlu dihitung terlebih dahulu turunan fungsi aktivasi tersebut. Turunan ReLU bernilai 1 jika input sebelum aktivasi lebih besar dari 0, dan bernilai 0 jika input kurang dari atau sama dengan 0, Setelah nilai error pada masing-masing *hidden layer* diperoleh, gradien terhadap bobot dan bias dihitung menggunakan persamaan 3.17 dan 3.18:

$$\frac{\partial L}{\partial W^{(l)}} = \delta^{(l)} A^{(l-1)} \tag{3.17}$$

$$\frac{\partial L}{\partial b^{(l)}} = \delta^{(l)} \tag{3.18}$$

Keterangan:

$\frac{\partial L}{\partial W^{(l)}}$: gradien *loss* terhadap bobot pada *layer* ke-l
$\frac{\partial L}{\partial b^{(l)}}$: gradien *loss* terhadap bias pada *layer* ke-l
$A^{(l-1)}$: *output* dari *layer* sebelumnya

Setelah mendapatkan error di lapisan *output*, kesalahan tersebut dirambatkan kembali (backpropagated) ke *Hidden Layer* 3. Sesuai Persamaan 3.15, error pada neuron tersembunyi (δ_3)sangat bergantung pada turunan fungsi aktivasi ReLU (Persamaan 3.16).

Karena keluaran *Hidden Layer* 3 sebelumnya bernilai positif ($H_1 = 0{,}60$dan $H_2 = 0{,}40$), maka turunan ReLU untuk kedua neuron tersebut bernilai 1. Error pada *Hidden Layer* 3 dihitung dengan mengalikan error *output* dengan bobot penghubungnya ($S_1 = 0{,}5$dan $S_2 = -0{,}3$):

a. Error pada Neuron 1 di *Hidden Layer* 3 $(\delta_{3,1})$

$$\delta_{3,1} = (S_1 \cdot \delta_{out}) \cdot f'(input_{H1})$$

$$\delta_{3,1} = (0{,}5 \cdot 0{,}57) \cdot 1 = 0{,}285$$

b. Error pada Neuron 2 di *Hidden Layer* 3 $(\delta_{3,2})$

$$\delta_{3,2} = (S_2 \cdot \delta_{out}) \cdot f'(input_{H2})$$

$$\delta_{3,2} = (-0{,}3 \cdot 0{,}57) \cdot 1 = -0{,}171$$

Dengan diketahuinya nilai error $\delta_3$tersebut, gradien bobot $(T_{i,m})$yang menghubungkan *Hidden Layer* 2 ke *Hidden Layer* 3 dapat dihitung dengan cara yang sama seperti Persamaan 3.17, yaitu mengalikan $\delta_3$dengan vektor keluaran dari *Hidden Layer* 2 (Z).

Proses perhitungan menggunakan aturan rantai (chain rule) ini terus diulang mundur hingga mencapai *Hidden Layer* 1, sehingga seluruh parameter di dalam

jaringan memiliki nilai gradiennya masing-masing ($\nabla_\theta L$)yang siap digunakan untuk tahapan optimasi. Melalui proses ini, diperoleh gradien fungsi *loss* terhadap seluruh parameter jaringan yang secara umum dinyatakan sebagai persamaan 3.19:

$$\nabla_\theta L \tag{3.19}$$

Keterangan:
$\nabla_\theta L$: gradien fungsi *loss* terhadap seluruh parameter jaringan
θ: kumpulan seluruh bobot dan bias dalam model

Setelah gradien pada masing-masing bobot dan bias dihitung, seluruh gradien tersebut dapat digabungkan menjadi satu kumpulan gradien parameter jaringan. Gradien inilah yang selanjutnya digunakan dalam tahap optimisasi untuk memperbarui parameter jaringan menggunakan algoritma Adam, sehingga model dapat secara iteratif meningkatkan kemampuan klasifikasinya terhadap data ekspresi gen normal dan kanker LUSC.

2. Optimasi Menggunakan Adam (*Adaptive Moment Estimation*)

Setelah proses *backpropagation* menghasilkan gradien fungsi *loss* terhadap seluruh parameter jaringan, tahap berikutnya adalah melakukan pembaruan bobot dan bias untuk meminimalkan nilai *loss*. Pada penelitian ini, proses optimisasi dilakukan menggunakan algoritma Adam (Adaptive Moment Estimation). Adam merupakan metode optimisasi berbasis gradien yang menggabungkan konsep Momentum dan RMSProp, sehingga mampu menyesuaikan laju pembelajaran dengan *learning rate* warmup (10 epoch pertama). Metode ini dipilih karena memiliki konvergensi yang cepat, stabil terhadap variasi skala gradien, serta efektif dalam menangani data berdimensi tinggi seperti data ekspresi gen. Pada setiap

iterasi ke-t, terlebih dahulu dihitung gradien fungsi *loss* terhadap parameter jaringan yang dinyatakan sebagaimana persamaan 3.20:

$$g_t = \nabla_\theta L_t \tag{3.20}$$

Keterangan:
g_t: gradien *loss* terhadap parameter pada iterasi ke-t
$\nabla_\theta L_t$: turunan fungsi *loss* terhadap parameter θ
θ: seluruh bobot dan bias dalam jaringan

Pada algoritma Adam, gradien terhadap parameter jaringan pada iterasi ke-t dinyatakan sebagai g_t. Nilai ini merupakan masukan awal untuk menghitung momen pertama dan momen kedua pada langkah optimisasi berikutnya. Gradien ini kemudian digunakan untuk menghitung estimasi momen pertama (first moment estimate), yaitu rata-rata eksponensial dari gradien, yang dirumuskan sebagaigaimana persamaan 3.21:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{3.21}$$

Keterangan:
m_t: estimasi momen pertama pada iterasi ke-t
m_{t-1}: estimasi momen pertama pada iterasi sebelumnya
β_1: koefisien peluruhan momen pertama
g_t: gradien pada iterasi ke-t

Selain itu, Adam juga menghitung estimasi momen kedua (second moment estimate), yaitu rata-rata eksponensial dari kuadrat gradien, sebagaimana persamaan 3.22:

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{3.22}$$

Keterangan:
v_t: estimasi momen kedua pada iterasi ke-t
v_{t-1}: estimasi momen kedua pada iterasi sebelumnya
β_2: koefisien peluruhan momen kedua
g_t^2: kuadrat gradien

Karena pada iterasi awal nilai $m_t$dan $v_t$cenderung bias terhadap nol, maka dilakukan koreksi bias menggunakan persamaan 3.23 dan 3.24:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{3.23}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{3.24}$$

Keterangan:
$\hat{m}_t$: momen pertama setelah koreksi bias
$\hat{v}_t$: momen kedua setelah koreksi bias
t: nomor iterasi pelatihan

Setelah diperoleh estimasi momen yang telah dikoreksi, parameter jaringan diperbarui menggunakan persamaan 3.25:

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{3.25}$$

Keterangan:
θ_t: parameter setelah pembaruan
θ_{t-1}: parameter sebelum pembaruan
α: *learning rate* (pada penelitian ini sebesar 0,001)
ϵ: konstanta kecil untuk menghindari pembagian dengan nol (umumnya 10^{-8})

Melalui mekanisme ini, pembaruan parameter dipengaruhi oleh rata-rata gradien dan variansinya secara simultan, sehingga langkah pembaruan menjadi lebih adaptif dan stabil dibandingkan metode gradient descent konvensional. Dengan penggunaan algoritma Adam, proses pelatihan model MLP pada data TCGA-LUSC diharapkan mampu mencapai konvergensi yang lebih cepat dan menghasilkan performa klasifikasi yang optimal dalam membedakan jaringan paru normal dan kanker LUSC.

Sebagai ilustrasi, akan dilakukan pembaruan parameter pada bobot S_1dari lapisan *output*. Berdasarkan perhitungan *backpropagation* sebelumnya, diperoleh nilai gradien $g_t = 0{,}342$. Nilai bobot lama sebelum diperbarui adalah $\theta_{t-1} = 0{,}5$.

Diasumsikan ini adalah iterasi pertama $(t = 1)$, sehingga nilai momen awal $m_0$dan $v_0$adalah 0, Konfigurasi bawaan algoritma Adam yang digunakan adalah: *learning rate* $\alpha = 0{,}001$; $\beta_1 = 0{,}9$; $\beta_2 = 0{,}999$; dan $\epsilon = 10^{-8}$.

a. Langkah 1: Menghitung Momen Pertama (m_t)

$$m_1 = \beta_1 m_0 + (1 - \beta_1) g_t$$

$$m_1 = 0{,}9(0) + (1 - 0{,}9)(0{,}342)$$

$$m_1 = 0{,}1 \cdot 0{,}342 = 0{,}0342$$

b. Langkah 2: Menghitung Momen Kedua (v_t)

$$v_1 = \beta_2 v_0 + (1 - \beta_2) g_t^2$$

$$v_1 = 0{,}999(0) + (1 - 0{,}999)(0{,}342)^2$$

$$v_1 = 0{,}001 \cdot 0{,}116964 = 0{,}000116964$$

c. Langkah 3: Koreksi Bias

$$m_1(\text{terkoreksi}) = \frac{0{,}0342}{1 - 0{,}9^1} = \frac{0{,}0342}{0{,}1} = 0{,}342$$

$$v_1(\text{terkoreksi}) = \frac{0{,}000116964}{1 - 0{,}999^1} = \frac{0{,}000116964}{0{,}001} = 0{,}116964$$

d. Langkah 4: Pembaruan Bobot

$$\theta_1 = \theta_{t-1} - \frac{\alpha \cdot m_1}{\sqrt{v_1} + \epsilon}$$

$$\theta_1 = 0{,}5 - \frac{0{,}001 \cdot 0{,}342}{\sqrt{0{,}116964} + 10^{-8}}$$

$$\theta_1 = 0{,}5 - \frac{0{,}000342}{0{,}342} = 0{,}5 - 0{,}001 = 0{,}499$$

Setelah proses optimasi, nilai bobot S_1yang semula 0,5 telah terkoreksi menjadi 0,499, Penurunan nilai bobot ini terjadi karena gradien bernilai positif, yang berarti peningkatan bobot justru akan meningkatkan nilai kesalahan (*loss*). Proses ini akan terus diulang untuk seluruh parameter jaringan pada setiap iterasi (epoch), sehingga model secara bertahap belajar membedakan pola ekspresi gen jaringan kanker LUSC dan jaringan normal

3.6 Prosedur Uji Coba

Bagian ini menguraikan prosedur uji coba untuk mengevaluasi performa model klasifikasi *Lung Squamous Cell Carcinoma* (LUSC) berbasis data ekspresi gen. Tahapan pengujian mencakup prapemrosesan data, seleksi fitur *Mutual Information* (MI), dan klasifikasi menggunakan arsitektur *Multilayer Perceptron* (MLP) dengan optimasi Adam. Guna menentukan konfigurasi model yang paling optimal, eksperimen dilakukan melalui tiga skenario utama, yaitu variasi jumlah *hidden layer*, jumlah neuron, dan jumlah fitur input. Performa dari setiap skenario tersebut kemudian dievaluasi secara kuantitatif berdasarkan *confusion matrix* untuk mengukur tingkat akurasi, presisi, *recall*, dan *F1-score*.

3.6.1 Skenario Penelitian

Penelitian ini dilakukan melalui beberapa skenario eksperimental yang dirancang untuk mengevaluasi pengaruh jumlah fitur hasil seleksi *Mutual Information* serta variasi arsitektur *Multilayer Perceptron* (MLP) terhadap performa klasifikasi ekspresi gen pada dataset TCGA-LUSC. Seluruh proses diawali dengan tahap *preprocessing* data, dilanjutkan dengan seleksi fitur

menggunakan Mutual Information, dan diakhiri dengan klasifikasi menggunakan MLP yang dioptimasi menggunakan algoritma Adam. Adapun skenario yang diterapkan dalam penelitian ini meliputi:

1. Skenario 1: Variasi *Hidden Layer*

   Skenario pertama merupakan studi ablasi yang bertujuan untuk menganalisis pengaruh jumlah *hidden layer* terhadap performa model *Multi Layer Perceptron* (MLP) dalam mempelajari pola non-linear pada data ekspresi gen. Variasi yang diuji pada skenario ini meliputi 2 dan 3 *hidden layer.*

   Melalui skenario ini, akan dianalisis pengaruh kedalaman arsitektur jaringan terhadap performa klasifikasi berdasarkan nilai akurasi, presisi, *recall*, dan F1-score.

2. Skenario 2: Variasi Neuron

   Skenario ini bertujuan untuk menentukan kapasitas jaringan yang paling optimal dalam merepresentasikan pola data tanpa menyebabkan underfitting maupun overfitting.

   Pada skenario ini, parameter lain seperti jumlah fitur input, dengan warmup learning rate sampai epoch ke 10, dan jumlah epoch maksimum tetap dipertahankan. Variasi jumlah neuron yang diuji meliputi beberapa konfigurasi arsitektur, yaitu:

   a. (64, 32, 16)
   b. (128, 64, 32)
   c. (256, 128, 64)
   d. (256,128)

e. (64,32)

f. (32,16)

Hasil dari pengujian pada skenario ini digunakan untuk menentukan konfigurasi jumlah neuron terbaik berdasarkan performa evaluasi model yang dihasilkan.

3. Skenario 3: Variasi Jumlah Fitur Input

Skenario terakhir bertujuan untuk menganalisis efektivitas reduksi dimensi terhadap performa klasifikasi. Mengingat karakteristik dataset yang bersifat *high-dimensional low-sample-size* (HDLSS), pemilihan jumlah fitur yang tepat sangat krusial untuk mencegah terjadinya overfitting. Oleh karena itu, dengan mengaplikasikan arsitektur jaringan dan learning rate terbaik yang telah diperoleh dari kedua tahap sebelumnya, pengujian pada skenario ini difokuskan pada empat variasi jumlah fitur input hasil seleksi Mutual Information (MI), yakni pengujian dengan menggunakan 50, 100, 500, dan 1500 fitur terbaik. Pendekatan pengujian secara bertahap ini memungkinkan pemilihan setiap parameter didasarkan pada bukti eksperimental yang kuat. Desain kombinasi eksperimen secara mendalam ditunjukkan pada Tabel 3.3 berikut.

Tabel 3. 3 Desain skenario pengujian

Split Data	**Model**	***Hidden Layer***	**Neuron *Hidden Layer***	**Jumlah Fitur Input**	***Max Iteration***
80:20	Model A	3	64, 32, 16	50	200
	Model B			100	
	Model C			500	
	Model D			1500	
	Model E	3	128, 64, 32	50	
	Model F			100	
	Model G			500	

Split Data	Model	*Hidden Layer*	Neuron *Hidden Layer*	Jumlah Fitur Input	*Max Iteration*
	Model H			1500	
	Model I	3	256, 128, 64	50	
	Model J			100	
	Model K			500	
	Model L			1500	
	Model M	2	32,16	50	
	Model N			100	
	Model O			500	
	Model P			1500	
	Model Q	2	128, 64	50	
	Model R			100	
	Model S			500	
	Model T			1500	
	Model U	2	256,128	50	
	Model V			100	
	Model W			500	
	Model X			1500	

3.6.2 Proses Pelatihan dan Pengujian Model

Pada tahap ini dilakukan proses pelatihan dan pengujian model klasifikasi ekspresi gen untuk membedakan jaringan normal dan jaringan kanker paru jenis *Lung Squamous Cell Carcinoma* (LUSC). Pendekatan yang digunakan adalah kombinasi seleksi fitur *Mutual Information* (MI) dan klasifikasi menggunakan *Multilayer Perceptron* (MLP) yang dioptimasi dengan algoritma Adam.

1. Pelatihan Model

Proses pelatihan model dilakukan secara iteratif mengikuti tiga tahapan skenario pengujian yang telah ditetapkan untuk menemukan konfigurasi parameter paling optimal . Tahap awal dimulai dengan pembagian dataset hasil *preprocessing* ke dalam data latih dan data uji menggunakan rasio 80% untuk data latih dan 20% untuk data uji yang dibagi secara *stratified* untuk menjaga proporsi distribusi kelas.

Pelatihan dijalankan dengan menyesuaikan parameter secara bertahap. Pada tahap pertama, dilakukan studi ablasi terhadap arsitektur MLP dengan memvariasikan jumlah lapisan tersembunyi (2 dan 3 layer) serta jumlah neuron per lapisan yang telah di tentukan untuk menentukan kapasitas jaringan yang paling efisien dalam menangani data ekspresi gen LUSC. Tahap terakhir adalah pengujian variasi jumlah fitur input hasil seleksi Mutual Information (50, 100, 500, dan 1500 fitur) menggunakan kombinasi arsitektur dan learning rate terbaik yang telah ditemukan sebelumnya .

Selama proses pelatihan di setiap skenario, data latih diproses melalui tahapan *feedforward propagation* untuk menghasilkan probabilitas prediksi kelas yang kemudian dihitung nilai kesalahannya menggunakan fungsi loss *Binary Cross-Entropy*. Setelah nilai loss diperoleh, dilakukan proses back*propagation* untuk menghitung gradien terhadap seluruh parameter jaringan, yaitu bobot dan bias pada setiap lapisan. Gradien tersebut digunakan oleh algoritma optimisasi Adam untuk memperbarui parameter model secara iteratif hingga mencapai 200 epoch atau hingga model menunjukkan kondisi konvergen. Dengan prosedur ini, model secara bertahap menyesuaikan parameter agar mampu meminimalkan kesalahan dan meningkatkan kemampuan klasifikasi terhadap data ekspresi gen.

2. Pengujian Model

Setelah proses pelatihan selesai, model diuji menggunakan data uji yang sebelumnya telah dipisahkan dari data latih. Pada tahap ini, parameter bobot dan bias hasil pelatihan tidak lagi diperbarui. Data uji terlebih dahulu direduksi menggunakan fitur yang sama seperti yang dipilih pada data latih. Kemudian data

tersebut diproses melalui jaringan MLP untuk menghasilkan probabilitas prediksi kelas. Untuk klasifikasi biner, probabilitas *output* dikonversi menjadi label kelas dengan aturan:

a. Jika $\hat{y} \geq 0{,}5$, maka diklasifikasikan sebagai kanker (1)
b. Jika $\hat{y} < 0{,}5$, maka diklasifikasikan sebagai normal (0)

Hasil prediksi kemudian dibandingkan dengan label aktual untuk menghitung performa model.

3. Confusion Matrix

Evaluasi performa dilakukan menggunakan *confusion matrix* yang terdiri dari empat komponen, yaitu: *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), *False Negative* (FN).

a. *True Positive* (TP): Jumlah sampel yang sebenarnya termasuk dalam kelas kanker LUSC (label 1) dan berhasil diprediksi dengan benar sebagai kanker oleh model.
b. *True Negative* (TN): Jumlah sampel yang sebenarnya termasuk dalam kelas normal (label 0) dan berhasil diprediksi dengan benar sebagai normal oleh model.
c. *False Positive* (FP): Jumlah sampel yang sebenarnya termasuk dalam kelas normal (label 0) dan berhasil diprediksi dengan benar sebagai normal oleh model.
d. *False Negative* (FN): Jumlah sampel yang sebenarnya termasuk dalam kelas kanker LUSC (label 1), namun diprediksi sebagai normal (label 0) oleh model. Kondisi ini disebut sebagai *Type II Error* dan dalam konteks medis

merupakan kesalahan yang sangat krusial karena kasus kanker tidak terdeteksi.

Tabel 3. 4 Hasil Deteksi Berdasarkan Kelas Positif dan Negatif

		Assignet Class	
Actual Class		*Positive*	*Negative*
	Positive	TP	FN
	Negative	FP	TN

Berdasarkan Tabel 3.4 yang menunjukkan hasil deteksi dalam bentuk confusion *matrix*, performa model selanjutnya dianalisis menggunakan beberapa metrik evaluasi kuantitatif. Keempat komponen pada confusion *matrix*, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN), digunakan sebagai dasar perhitungan untuk mengukur tingkat akurasi dan kemampuan model dalam mengklasifikasikan data secara tepat. Metrik evaluasi tersebut meliputi akurasi, presisi, *recall*, dan F1-score yang dirumuskan sebagai berikut.

a. Akurasi

Akurasi menunjukkan persentase prediksi yang benar terhadap seluruh data uji, dirumuskan sebagai:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Akurasi = \frac{45 + 40}{45 + 40 + 5 + 10}$$

$$Akurasi = \frac{85}{100} = 0{,}85$$

b. Presisi

Presisi menunjukkan proporsi prediksi positif yang benar-benar merupakan kelas positif, dirumuskan sebagai:

$$Presisi = \frac{TP}{TP + FP}$$

$$Presisi = \frac{45}{45 + 5}$$

$$Presisi = \frac{45}{50} = 0{,}90$$

c. *Recall* (Sensitivitas)

Recall menunjukkan kemampuan model dalam mendeteksi seluruh kasus positif yang sebenarnya, dirumuskan sebagai:

$$Recall = \frac{TP}{TP + FN}$$

$$Recall = \frac{45}{45 + 10}$$

$$Recall = \frac{45}{55} \approx 0{,}818$$

d. F1-Score

F1-Score merupakan rata-rata harmonik antara presisi dan *recall*, dirumuskan sebagai:

$$F1\text{-}Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall}$$

$$F1\text{-}Score = 2 \times \frac{0{,}90 \times 0{,}818}{0{,}90 + 0{,}818}$$

$$F1\text{-}Score = 2 \times \frac{0{,}7362}{1.718} \approx 0{,}857$$

Melalui metrik-metrik tersebut, performa setiap kombinasi skenario (variasi jumlah fitur dan jumlah *hidden layer*) dibandingkan untuk menentukan konfigurasi terbaik dalam mengklasifikasikan data ekspresi gen TCGA-LUSC.

3.7 Implementasi Sistem

Sub-bab ini menguraikan tahapan implementasi komputasional dari desain sistem yang telah dirancang. Implementasi ini direpresentasikan dalam bentuk *Pseudocode* algoritmik yang berfokus pada logika dasar murni matematika dan struktur data, tanpa bergantung pada sintaks pustaka (*library*) pemrograman spesifik. Representasi algoritmik ini dibagi ke dalam empat modul utama: prapemrosesan data, inisialisasi parameter jaringan, forward *propagation*, serta *backpropagation* yang dioptimasi menggunakan algoritma Adam.

3.7.1 Modul Prapemrosesan Data Ekspresi Gen

Tahapan prapemrosesan mencakup transformasi skala data (*scaling*) menggunakan metode *Standard Scaler* agar setiap fitur gen memiliki distribusi dengan rata-rata mendekati nol dan standar deviasi sebesar satu, sesuai dengan Persamaan 3.2.

***Pseudocode* 3.1** Modul Prapemrosesan Data (*Standardization*)

```
FUNCTION StandardScaler(Dataset):
  // Dataset adalah matriks berukuran N (sampel) x F (fitur)
  N = JUMLAH_BARIS(Dataset)
  F = JUMLAH_KOLOM(Dataset)
  Dataset_Scaled = BUAT_MATRIKS_KOSONG(N, F)

  FOR j = 1 TO F DO
    // Hitung rata-rata (mean) untuk fitur ke-j
    Sum = 0
    FOR i = 1 TO N DO
      Sum = Sum + Dataset[i][j]
    END FOR
    Mean_j = Sum / N

    // Hitung varians dan standar deviasi untuk fitur ke-j
    Sum_Variance = 0
    FOR i = 1 TO N DO
      Sum_Variance = Sum_Variance + (Dataset[i][j] - Mean_j)^2
    END FOR
    Std_j = AKAR_KUADRAT(Sum_Variance / N)

    // Terapkan standarisasi Z-Score
```

```
    FOR i = 1 TO N DO
      // Mencegah pembagian dengan nol (zero division error)
      IF Std_j == 0 THEN
        Dataset_Scaled[i][j] = 0
      ELSE
        Dataset_Scaled[i][j] = (Dataset[i][j] - Mean_j) / Std_j
      END IF
    END FOR
  END FOR

  RETURN Dataset_Scaled
END FUNCTION
```

Algoritma pada *Pseudocode* 3.1 menerima masukan berupa matriks data ekspresi gen hasil augmentasi SMOTE. Proses iterasi pertama dilakukan secara vertikal (per kolom fitur) untuk menghitung nilai sentralitas (rata-rata) dan dispersi (standar deviasi) dari masing-masing gen. Selanjutnya, algoritma mentransformasikan setiap nilai sampel individual dengan mengurangkan nilai rata-rata dan membaginya dengan standar deviasi. Kondisional bersyarat (IF Std_j == 0) ditambahkan sebagai mekanisme penanganan error apabila terdapat gen yang ekspresinya konstan pada seluruh sampel, sehingga menjaga stabilitas komputasi sebelum data diumpankan ke dalam model *Multilayer Perceptron* (MLP).

3.7.2 Modul Seleksi Fitur dengan Mutual Information

Data yang telah distandarisasi kemudian direduksi dimensinya untuk mengatasi permasalahan *high-dimensional low-sample-size* (HDLSS). Proses komputasi *Mutual Information* (MI) mengukur tingkat ketergantungan probabilitas antara masing-masing fitur genetik terhadap label kelas (Persamaan 3.3).

***Pseudocode* 3.2** Modul Seleksi Fitur (Mutual Information)

```
FUNCTION MutualInformationSelection(Dataset_X, Labels_Y, K_Feature s):
  // Dataset_X: Matriks fitur N x F (setelah standarisasi)
  // Labels_Y: Vektor target kelas (0 atau 1)
```

```
  // K_Feature s: Jumlah batas fitur terbaik yang ingin diekstrak
(misal: 100)

  N = JUMLAH_SAMPEL(Dataset_X)
  F = JUMLAH_FITUR(Dataset_X)
  MI_Scores = LIST_KOSONG(F)

  // 1. Hitung probabilitas marginal kelas Y
  P_Y = HITUNG_PROBABILITAS_KELAS(Labels_Y)

  FOR j = 1 TO F DO
    Feature _j = Dataset_X[:, j]

    // 2. Diskretisasi fitur kontinu ke dalam n-bin untuk estimasi PDF
(Probability Density Function)
    Bins_X = DISKRETISASI_N_BIN(Feature _j, Bins = 10)

    // 3. Hitung probabilitas marginal P(X) dan probabilitas gabungan
P(X, Y)
    P_X = HITUNG_PROBABILITAS_MARGINAL(Bins_X)
    P_XY = HITUNG_PROBABILITAS_GABUNGAN(Bins_X, Labels_Y)

    MI_Value = 0

    // 4. Kalkulasi nilai Information Gain / Mutual Information
    FOR SETIAP Kelas y DALAM Labels_Y DO
      FOR SETIAP Bin x DALAM Bins_X DO
        IF P_XY[x, y] > 0 THEN
          // Sesuai Persamaan 3.3: P(x,y) * log( P(x,y) / (P(x)*P(y))
)
          Rasio = P_XY[x, y] / (P_X[x] * P_Y[y])
          MI_Value = MI_Value + P_XY[x, y] * LOGARITMA_NATURAL(Rasio)
        END IF
      END FOR
    END FOR

    MI_Scores[j] = MI_Value
  END FOR

  // Urutkan fitur berdasarkan skor MI secara menurun (descending)
  Sorted_Indices = URUTKAN_INDEKS_MENURUN(MI_Scores)

  // Pilih indeks gen sebanyak K terbaik
  Top_K_Indices = AMBIL_ELEMEN(Sorted_Indices, 1 SAMPAI K_Feature s)

  // Ekstrak dataset baru hanya dengan fitur yang terpilih
  Dataset_Reduced = EKSTRAK_KOLOM(Dataset_X, Top_K_Indices)

  RETURN Dataset_Reduced, Top_K_Indices
END FUNCTION
```

Algoritma pada *Pseudocode* 3.2 menjabarkan tahapan ekstraksi fitur untuk menemukan K-gen yang paling diskriminatif. Mengingat nilai ekspresi gen merupakan variabel kontinu, algoritma melakukan pendekatan fungsi kepadatan probabilitas dengan cara mendiskretisasi rentang nilai fitur ke dalam beberapa interval (bins). Algoritma menghitung probabilitas marginal kemunculan setiap bin $P(X)$, probabilitas kelas target $P(Y)$, serta probabilitas gabungan $P(X,Y)$. Nilai *Mutual Information* dihitung dengan mengakumulasikan perkalian antara probabilitas gabungan dengan logaritma natural dari rasio ketergantungannya (sesuai Persamaan 3.3). *Array* skor tersebut kemudian diurutkan secara menurun (*descending*), dan matriks dataset yang baru dibentuk hanya menyertakan vektor gen dengan skor informasi paling tinggi, membuang sisa fitur yang redundan.

3.7.3 Modul Inisialisasi Arsitektur MLP dan Parameter Adam

Sebelum proses pelatihan dimulai, arsitektur jaringan saraf tiruan perlu dibangun melalui inisialisasi parameter bobot (W) dan bias (b). Selain itu, variabel memori untuk algoritma optimasi Adam (momen pertama dan kedua) juga diinisialisasi pada tahap ini.

***Pseudocode* 3.3** Modul Inisialisasi Parameter

```
FUNCTION InitializeMLP(Input_Size, Hidden _Sizes, Output_Size):
  // Input_Size: Jumlah fitur K hasil seleksi Mutual Information
  // Hidden _Sizes: Array berisi jumlah neuron di tiap hidden layer
(misal: [64, 32, 16])

  Weights = LIST_KOSONG()
  Biases = LIST_KOSONG()
  Adam_M = LIST_KOSONG() // Estimasi momen pertama
  Adam_V = LIST_KOSONG() // Estimasi momen kedua

  Layer_Sizes = GABUNGKAN([Input_Size], Hidden _Sizes, [Output_Size])
  Total_Layers = PANJANG(Layer_Sizes)

  FOR L = 1 TO Total_Layers - 1 DO
```

```
    Node_In = Layer_Sizes[L]
    Node_Out = Layer_Sizes[L+1]

    // Inisialisasi bobot dengan nilai acak berdistribusi normal kecil
(He/Xavier Initialization)
    W_L = BANGKITKAN_MATRIKS_ACAK(Node_In, Node_Out) * 0,01
    b_L = BANGKITKAN_VEKTOR_NOL(Node_Out)

    // Inisialisasi matriks memori Adam dengan nilai nol
    M_W = BANGKITKAN_MATRIKS_NOL(Node_In, Node_Out)
    V_W = BANGKITKAN_MATRIKS_NOL(Node_In, Node_Out)
    M_b = BANGKITKAN_VEKTOR_NOL(Node_Out)
    V_b = BANGKITKAN_VEKTOR_NOL(Node_Out)

    TAMBAHKAN(Weights, W_L)
    TAMBAHKAN(Biases, b_L)
    TAMBAHKAN(Adam_M, {W: M_W, b: M_b})
    TAMBAHKAN(Adam_V, {W: V_W, b: V_b})
  END FOR

  RETURN Weights, Biases, Adam_M, Adam_V
END FUNCTION
```

Pseudocode 3.3 bertugas untuk membangun topologi jaringan secara dinamis berdasarkan parameter dimensi yang diberikan. Matriks bobot antar lapisan diinisialisasi menggunakan nilai acak skalar kecil untuk memecah simetri (*symmetry breaking*), yang memungkinkan setiap neuron mempelajari fitur yang berbeda selama pelatihan. Variabel Adam_M dan Adam_V merepresentasikan *first moment* dan *second moment* dari algoritma Adam, yang diinisialisasi dengan matriks/vektor bernilai nol murni sesuai dengan kaidah matematis pada sistem dinamis stokastik. Seluruh struktur data ini disimpan ke dalam bentuk *list* (kumpulan array) untuk mempermudah pemanggilan rekursif pada saat komputasi *forward* dan *backward*.

3.7.4 Modul Forward *Propagation*

Modul ini merepresentasikan aliran data dari lapisan input, melewati serangkaian lapisan tersembunyi (Persamaan 3.5 - 3.7), hingga menghasilkan probabilitas luaran di lapisan *output* (Persamaan 3.8).

***Pseudocode* 3.4** Modul *Forward Propagation*

```
FUNCTION ForwardPropagation(X_Batch, Weights, Biases):
  Activations = LIST_KOSONG()
  Z_Values = LIST_KOSONG()

  A_Current = X_Batch
  TAMBAHKAN(Activations, A_Current)

  Total_Hidden _Layers = PANJANG(Weights) - 1

  // Aliran komputasi untuk Hidden Layers
  FOR L = 1 TO Total_Hidden _Layers DO
    // Transformasi Linear: Z = A_prev * W + b
    Z_Current = PERKALIAN_MATRIKS(A_Current, Weights[L]) + Biases[L]
    TAMBAHKAN(Z_Values, Z_Current)

    // Fungsi Aktivasi ReLU: max(0, Z)
    A_Current = MAX(0, Z_Current)
    TAMBAHKAN(Activations, A_Current)
  END FOR

  // Aliran komputasi untuk Output Layer
  Z_Out = PERKALIAN_MATRIKS(A_Current, Weights[Total_Hidden _Layers +
1]) + Biases[Total_Hidden _Layers + 1]
  TAMBAHKAN(Z_Values, Z_Out)

  // Fungsi Aktivasi Sigmoid
  Y_Hat = 1 / (1 + EKSPONENSIAL(-Z_Out))
  TAMBAHKAN(Activations, Y_Hat)

  RETURN Y_Hat, Activations, Z_Values
END FUNCTION
```

Proses *feedforward* berjalan dengan memproses sekumpulan data masukan (X_Batch) secara berurutan. Algoritma melakukan operasi aljabar linear berupa perkalian titik (*dot product*) antara matriks aktivasi dari lapisan sebelumnya dengan matriks bobot, lalu ditambahkan dengan vektor bias (*Z_Current*). Untuk lapisan

tersembunyi, matriks Z ditransformasikan menggunakan fungsi non-linear ReLU yang akan mengeliminasi nilai negatif menjadi nol, sehingga mengatasi masalah *vanishing gradient*. Pada iterasi terakhir, lapisan keluaran memproses data menggunakan fungsi aktivasi Sigmoid untuk memetakan ruang nilai kontinu murni ke dalam rentang probabilitas [0, 1] yang mengindikasikan kelas LUSC. Seluruh luaran *intermediate* (*Activations* dan *Z_Values*) disimpan dalam memori (*cache*) karena nilai-nilai ini mutlak dibutuhkan untuk perhitungan rantai turunan pada saat *backpropagation*.

3.7.5 Modul *Backpropagation* dan Pembaruan Parameter (*Adam Optimizer*)

Modul ini adalah inti dari proses pembelajaran model. Di sini, kesalahan (*error*) dievaluasi dan dikalkulasikan mundur untuk menghasilkan gradien parameter (Persamaan 3.12 - 3.18). Selanjutnya, parameter jaringan diperbarui menggunakan estimasi momen dari algoritma Adam (Persamaan 3.20 - 3.25).

***Pseudocode* 3.5** Modul *Backpropagation* dan Pembaruan Parameter (*Adam Optimizer*)

```
FUNCTION BackpropAndUpdate(Y_True , Y_Hat, Activations, Z_Values, Weights, Biases, Adam_M, Adam_V, t_Iter, Alpha):
  // Konstanta Adam
  Beta1 = 0,9; Beta2 = 0,999; Epsilon = 10^-8
  M = JUMLAH_SAMPEL(Y_True ) // Ukuran batch

  Gradients_W = LIST_KOSONG_SEUKURAN(Weights)
  Gradients_b = LIST_KOSONG_SEUKURAN(Biases)

  Total_Layers = PANJANG(Weights)

  // 1. Hitung Error di Lapisan Output (Turunan BCE & Sigmoid)
  // dZ_out = Y_Hat - Y_True
  dZ = Y_Hat - Y_True

  // 2. Loop Backpropagation dari lapisan terakhir menuju lapisan pertama
  FOR L = Total_Layers DOWNTO 1 DO
    A_Prev = Activations[L - 1]

    // Kalkulasi Gradien Bobot dan Bias
    dW = PERKALIAN_MATRIKS(TRANSPOSE(A_Prev), dZ) / M
```

```
    db = RATA_RATA_KOLOM(dZ)

    Gradients_W[L] = dW
    Gradients_b[L] = db

    // Kalkulasi error untuk layer sebelumnya (jika bukan input layer)
    IF L > 1 THEN
      W_Current = Weights[L]
      Z_Prev = Z_Values[L - 1]

      // Turunan aturan rantai (Chain Rule)
      dA_Prev = PERKALIAN_MATRIKS(dZ, TRANSPOSE(W_Current))

      // Turunan fungsi aktivasi ReLU
      dReLU = BENTUK_MATRIKS(Z_Prev > 0 ? 1 : 0)
      dZ = dA_Prev * dReLU // Perkalian elemen (Hadamard product)
    END IF
  END FOR

  // 3. Pembaruan Parameter menggunakan Algoritma Adam
  FOR L = 1 TO Total_Layers DO
    // Hitung estimasi momen untuk Bobot (W)
    Adam_M[L].W = (Beta1 * Adam_M[L].W) + ((1 - Beta1) * Gradients_W[L])
    Adam_V[L].W = (Beta2 * Adam_V[L].W) + ((1 - Beta2) *
(Gradients_W[L]^2))

    // Hitung estimasi momen untuk Bias (b)
    Adam_M[L].b = (Beta1 * Adam_M[L].b) + ((1 - Beta1) * Gradients_b[L])
    Adam_V[L].b = (Beta2 * Adam_V[L].b) + ((1 - Beta2) *
(Gradients_b[L]^2))

    // Koreksi Bias (Bias Correction)
    M_W_hat = Adam_M[L].W / (1 - Beta1^t_Iter)
    V_W_hat = Adam_V[L].W / (1 - Beta2^t_Iter)
    M_b_hat = Adam_M[L].b / (1 - Beta1^t_Iter)
    V_b_hat = Adam_V[L].b / (1 - Beta2^t_Iter)

    // Perbarui Bobot dan Bias model
    Weights[L] = Weights[L] - (Alpha * M_W_hat) / (AKAR_KUADRAT(V_W_hat)
+ Epsilon)
    Biases[L] = Biases[L] - (Alpha * M_b_hat) / (AKAR_KUADRAT(V_b_hat)
+ Epsilon)
  END FOR

  RETURN Weights, Biases, Adam_M, Adam_V
END FUNCTION

  RETURN Y_Hat, Activations, Z_Values
END FUNCTION
```

Proses *backpropagation* dimulai dengan menghitung derivatif fungsi objektif *Binary Cross-Entropy* terhadap probabilitas prediksi keluaran (vektor dZ). Dengan memanfaatkan kalkulus aturan rantai (*chain rule*), gradien disebarkan ke arah mundur secara iteratif. Pada lapisan tersembunyi, gradien dikalikan dengan turunan fungsi ReLU ($dReLU$), yang bertindak sebagai gerbang (*gate*); meneruskan gradien bernilai utuh untuk nilai input positif, dan menganulir gradien (menjadikannya nol) untuk input negatif atau nol.

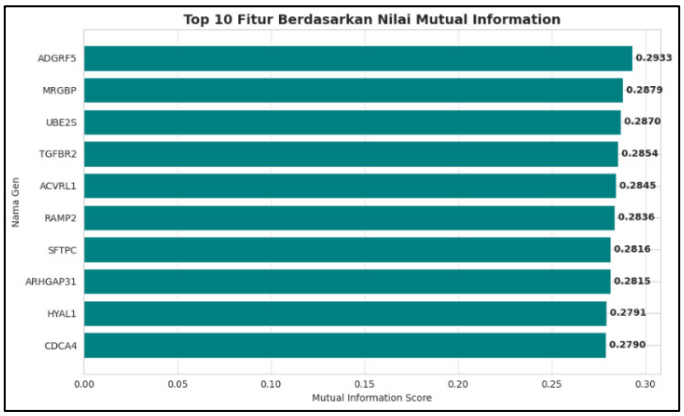
Setelah seluruh gradien parameter (dW dan db) diperoleh, algoritma Adam mengambil alih proses optimisasi. Gradien tersebut tidak serta-merta digunakan untuk mengurangi bobot secara linear, melainkan diproses untuk memutakhirkan memori momen eksponensial pertama (Adam_M, yang merepresentasikan arah momentum) dan momen eksponensial kedua tak-terpusat (Adam_V, yang merepresentasikan varians gradien). Tahap *bias correction* diterapkan untuk mencegah nilai estimasi menuju nol pada masa iterasi awal (t_Iter). Akhirnya, parameter bobot dan bias jaringan disesuaikan (diperbarui) dengan mengurangkan nilainya berdasarkan skala dinamis *learning rate* (*Alpha*) yang adaptif untuk setiap parameter genetik yang dilibatkan.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil Pengujian

Pengujian dalam penelitian ini diawali dengan tahap persiapan data menggunakan dataset TCGA-LUSC yang terdiri dari 561 sampel, meliputi 49 sampel jaringan normal dan 512 sampel jaringan kanker. Data ekspresi gen tersebut melalui serangkaian proses pra-pemrosesan, dimulai dari *encoding* label kelas, normalisasi rentang nilai menggunakan *StandardScaler*, hingga penanganan ketidakseimbangan kelas (*class imbalance*) pada data latih menggunakan teknik SMOTE (*Synthetic Minority Over-sampling Technique*). Setelah itu, reduksi dimensi dan pemilihan fitur dilakukan melalui metode seleksi fitur berbasis *Mutual Information* (MI) untuk menyaring fitur-fitur genetik yang memiliki tingkat relevansi tertinggi terhadap target klasifikasi kanker paru. Sepuluh fitur gen teratas dengan nilai *Mutual Information* tertinggi dari proses seleksi ini dapat dilihat pada Gambar 4.1.



Gambar 4. 1 Top 10 Fitur Berdasarkan Nilai Mutual Information

Setelah fitur-fitur terbaik diekstraksi, tahap selanjutnya adalah pembentukan dan konfigurasi arsitektur *Multilayer Perceptron* (MLP). Berdasarkan skenario pengujian yang telah dirancang pada Tabel 3.3, sebanyak dua puluh empat konfigurasi model MLP dibangun untuk diuji kinerjanya. Pengujian konfigurasi ini dilakukan dengan memvariasikan tiga *hyperparameter* utama, yaitu jumlah *hidden layer*, variasi jumlah neuron pada setiap *layer*, dan jumlah fitur input *(K-Best)* yang digunakan.

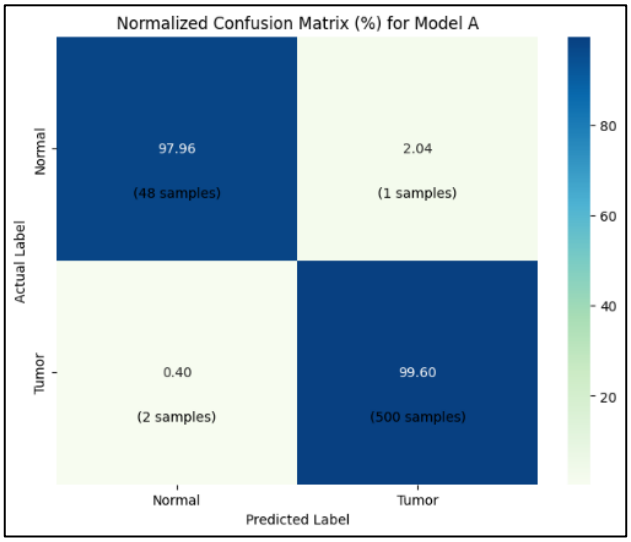
Pada tahap akhir, setiap model dievaluasi secara komprehensif menggunakan metode *5-Fold Cross validation*, di mana pada setiap iterasi (*fold*), data didistribusikan secara proporsional menjadi 80% data latih dan 20% data uji. Kinerja prediksi dari masing-masing arsitektur dievaluasi berdasarkan matriks kebingungan (*confusion matrix*) guna memperoleh nilai rata-rata dari metrik akurasi, presisi, *recall*, dan *F1-score*. Di samping perhitungan nilai rata-rata tersebut, standar deviasi tingkat akurasi antar *fold* juga turut dicatat sebagai indikator empiris untuk mengukur tingkat kestabilan dan keandalan model secara keseluruhan.

4.1.1 Model A

Model A dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik.

Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.



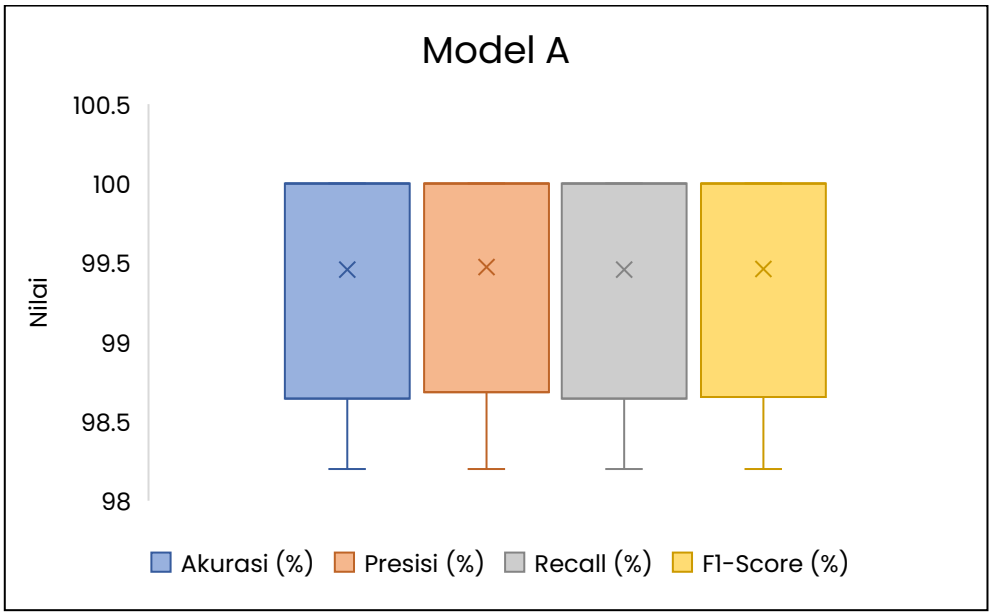
Gambar 4. 2 Metrik Evaluasi Model A

Tabel 4. 1 Metrik Evaluasi Model A

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,20	98,20	98,20	98,20
Fold 2	100	100	100	100
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100	100	100	100
Fold 5	100	100	100	100
Rata-rata	**99,46**	**99,47**	**99,460**	**99,46**

Untuk kelas Normal, model berhasil mengenali 48 sampel secara benar (*true negative*). Adanya 1 *false negative* mengindikasikan bahwa sebagian kecil profil ekspresi gen pada sampel Tumor memiliki karakteristik yang menyerupai jaringan Normal, sehingga menyulitkan pemisahan batas keputusan. Hal ini merupakan tantangan inheren pada klasifikasi berbasis data genomik berdimensi tinggi. Meskipun demikian, nilai rata-rata F1-Score keseluruhan sebesar 99,46% mengonfirmasi bahwa kedua kelas dapat dikenali dengan tingkat akurasi yang

sangat tinggi, dan kelas Tumor secara umum lebih mudah diidentifikasi berkat pola ekspresi gen yang lebih menonjol pada sampel kanker LUSC.



Gambar 4. 3 Visualisasi Rata-rata evaluation *matrix* Model A

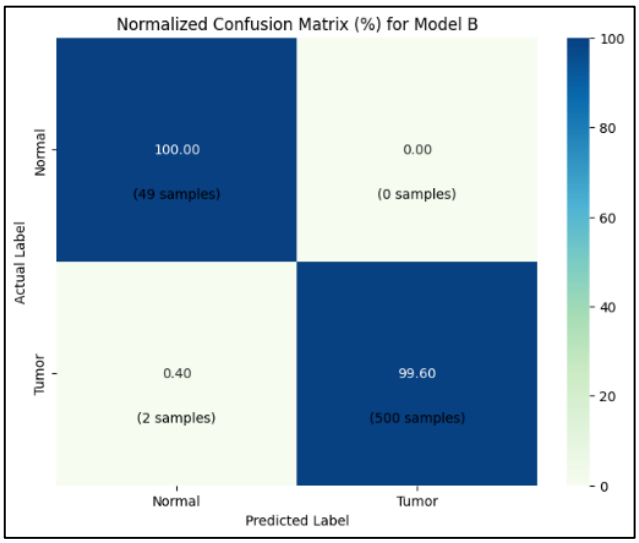
Lebih lanjut, distribusi performa Model A divisualisasikan melalui grafik *boxplot* pada Gambar 4.3. Berdasarkan grafik tersebut, diperoleh nilai rata-rata (μ) dan standar deviasi (σ) untuk Akurasi μ=0,9945 (σ=0,0073), Presisi (*Macro*) μ=0,9799 (σ=0,0248), *Recall* (*Macro*) μ=0,9880 (σ=0,0216), dan F1 (*Macro*) μ=0,9837 (σ=0,0219). Secara visual, kotak IQR Akurasi tampak relatif sempit di rentang nilai tinggi, sedangkan kotak IQR Presisi (*Macro*) dan F1 (*Macro*) jauh lebih lebar dengan whisker yang panjang ke bawah, mencerminkan adanya variasi anta*fold* yang cukup besar pada kedua metrik tersebut. Selain itu, terdapat sebuah titik pencilan pada metrik *Recall* (*Macro*) yang menunjukkan adanya satu *fold* dengan performa yang menyimpang secara signifikan. Kondisi ini mengindikasikan

bahwa Model A cukup stabil pada metrik Akurasi, namun masih mengalami fluktuasi performa yang perlu diperhatikan pada dimensi Presisi dan F1.

4.1.2 Model B

Model B dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru.



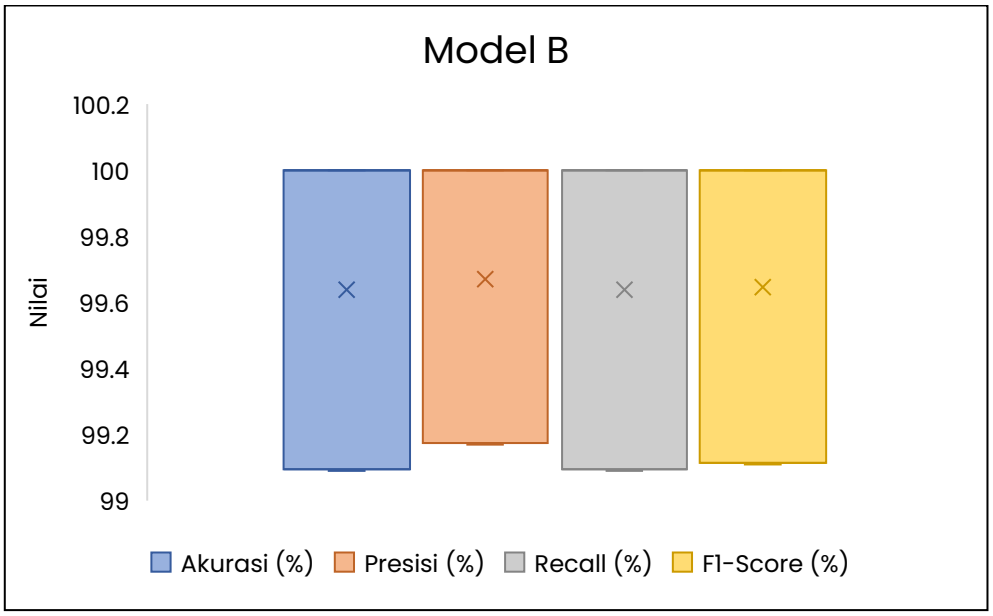
Gambar 4. 4 Metrik Evaluasi Model B

Tabel 4. 2 Metrik Evaluasi Model B

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.4, Model B mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,67%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 5 Visualisasi Rata-rata evaluation *matrix* Model B

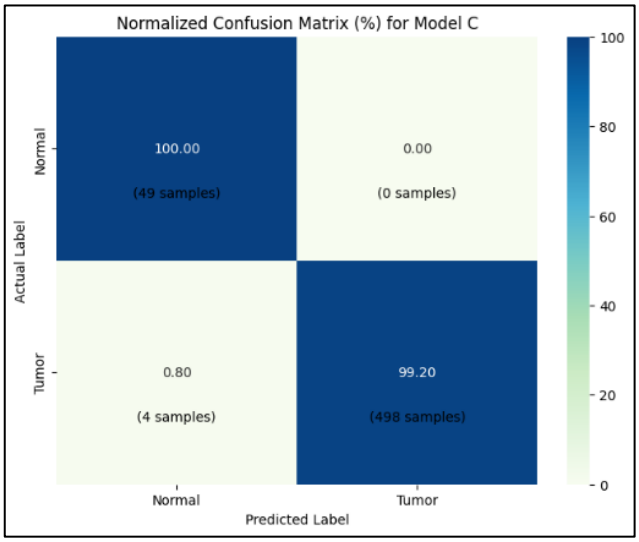
Lebih lanjut, distribusi performa Model B divisualisasikan melalui grafik *boxplot* pada Gambar 4.5. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Kotak IQR pada Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi, dengan σ *Recall*

sebesar 0,0024 yang merupakan nilai sangat kecil, mencerminkan kestabilan yang sangat tinggi pada kedua metrik tersebut. Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah yang panjang, mencerminkan variasi yang substansial pada dimensi presisi di antara *fold* pengujian. Secara keseluruhan, Model B menunjukkan konsistensi yang sangat baik pada Akurasi dan *Recall*, namun fluktuasi pada Presisi (*Macro*) masih perlu diperhatikan.

4.1.3 Model C

Model C dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

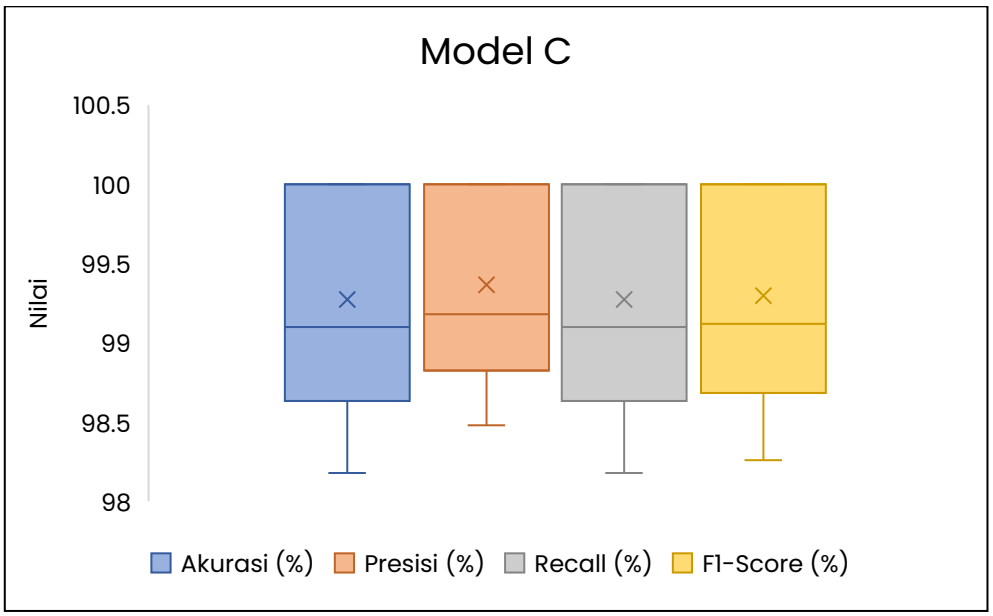


Gambar 4. 6 Metrik Evaluasi Model C

Tabel 4. 3 Metrik Evaluasi Model C

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	98,18	98,48	98,18	98,26
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,27**	**99,37**	**99,27**	**99,30**

Berdasarkan Gambar 4.6, Model C mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 497 data kanker sebagai *true positive*, dengan 4 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,27%, presisi 99,37%, *recall* 99,27%, dan F1-score 99,30%. Nilai standar deviasi sebesar 0,0068 menunjukkan konsistensi performa yang cukup stabil.



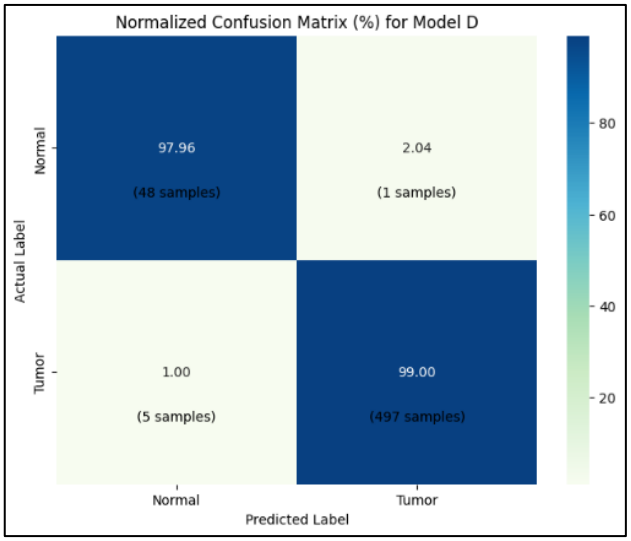
Gambar 4. 7 Visualisasi Rata-rata evaluation *matrix* Model C

Lebih lanjut, distribusi performa Model C divisualisasikan melalui grafik *boxplot* pada Gambar 4.7. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9927 (σ=0,0068), Presisi (*Macro*) μ=0,9652 (σ=0,0316), *Recall* (*Macro*) μ=0,9960 (σ=0,0037), dan F1 (*Macro*) μ=0,9794 (σ=0,0190). Kotak IQR *Recall* (*Macro*) tampak sangat tipis dengan whisker yang pendek, sesuai dengan nilai σ=0,0037 yang kecil, mencerminkan kestabilan yang sangat tinggi pada metrik tersebut. Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar dengan whisker bawah yang menjangkau nilai sangat rendah, sebagaimana dikonfirmasi oleh nilai σ=0,0316 yang merupakan terbesar di antara seluruh metrik Model C. Kondisi ini mengindikasikan bahwa Model C sangat konsisten pada *Recall*, namun mengalami fluktuasi yang cukup signifikan pada dimensi Presisi di antara partisi data yang berbeda.

4.1.4 Model D

Model D dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

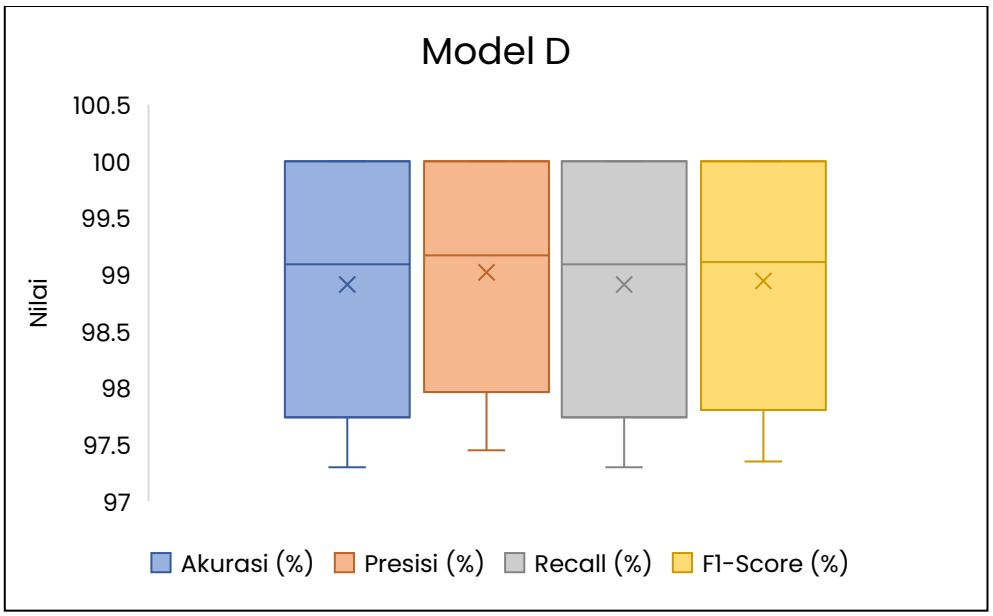


Gambar 4. 8 Metrik Evaluasi Model D

Tabel 4. 4 Metrik Evaluasi Model D

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	97.30	97.45	97.30	97.35
Fold 2	99,09	99,17	99,09	99,11
Fold 3	98,18	98,48	98,18	98,26
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**98,91**	**99,02**	**98,91**	**98,94**

Berdasarkan Gambar 4.8, Model D mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 496 data kanker sebagai *true positive*, dengan 1 *false negative* dan 5 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 98,91%, presisi 98,98%, *recall* 98,91%, dan F1-score 98,93%. Nilai standar deviasi sebesar 0,0105 merupakan yang tertinggi di antara seluruh model, mengindikasikan fluktuasi performa yang signifikan antar *fold* pengujian.



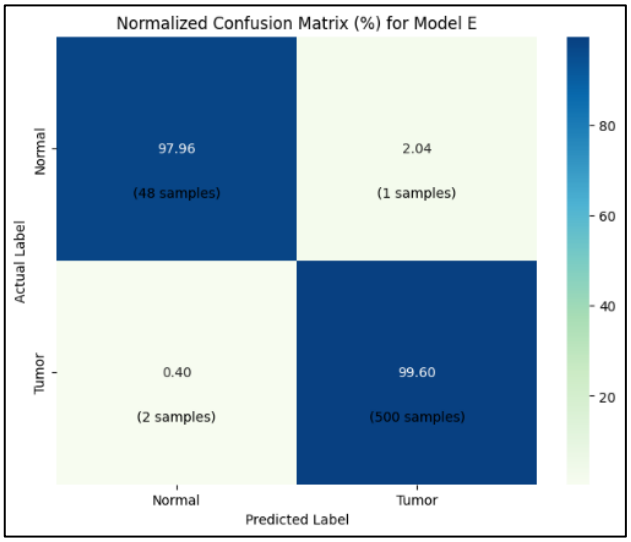
Gambar 4. 9 Visualisasi Rata-rata evaluation *matrix* Model D

Lebih lanjut, distribusi performa Model D divisualisasikan melalui grafik *boxplot* pada Gambar 4.9. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9891 (σ=0,0106), Presisi (*Macro*) μ=0,9551 (σ=0,0403), *Recall* (*Macro*) μ=0,9850 (σ=0,0228), dan F1 (*Macro*) μ=0,9688 (σ=0,0304). Secara visual, kotak IQR pada seluruh metrik tampak lebih lebar dibandingkan sebagian besar model lainnya, dengan whisker yang panjang ke berbagai arah. Pada metrik *Recall* (*Macro*), terdapat sebuah titik pencilan yang berada jauh di bawah batas whisker, mengindikasikan adanya satu *fold* dengan performa yang sangat menyimpang. Nilai

σ Presisi (*Macro*) sebesar 0,0403 dan σ Akurasi sebesar 0,0106 yang keduanya tergolong tinggi mengindikasikan bahwa Model D merupakan model dengan tingkat fluktuasi performa yang paling besar di antara seluruh skenario, sehingga keandalannya pada partisi data yang berbeda perlu dipertimbangkan secara hati-hati.

4.1.5 Model E

Model E diuji menggunakan pembagian data 80:20 yang menghasilkan 440 data latih dan 121 data uji. Arsitektur yang digunakan terdiri atas tiga *hidden layer* berukuran (64, 32, 16), menggunakan 100 fitur, dengan batch size 32, dan hasil pengujiannya ditampilkan pada Gambar 4.1 serta dirangkum pada Tabel 4.1.

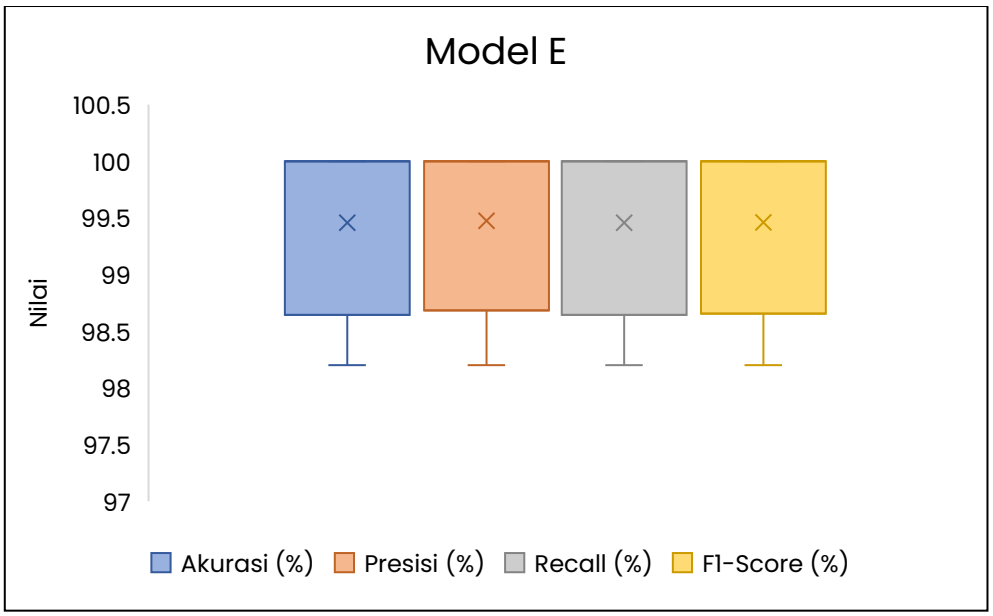


Gambar 4. 10 Metrik Evaluasi Model E

Tabel 4. 5 Metrik Evaluasi Model E

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,20	98,20	98,20	98,20
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**99,47**	**99,46**	**99,46**

Berdasarkan Gambar 4.10, Model E mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 499 data kanker sebagai *true positive*, dengan 1 *false negative* dan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 99,46%, *recall* 99,46%, dan F1-score 99,46%. Nilai standar deviasi sebesar 0,0072 menunjukkan adanya fluktuasi performa yang perlu diperhatikan di antara *fold* pengujian.



Gambar 4. 11 Visualisasi Rata-rata evaluation *matrix* Model E

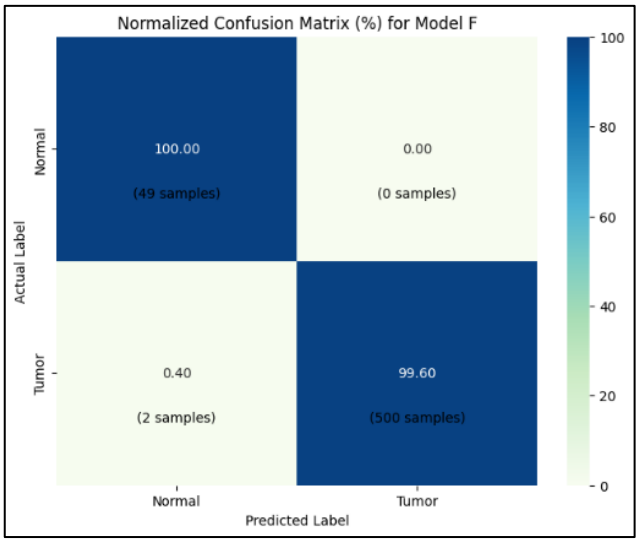
Lebih lanjut, distribusi performa Model E divisualisasikan melalui grafik *boxplot* pada Gambar 4.11. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0073), Presisi (*Macro*) μ=0,9799 (σ=0,0248), *Recall* (*Macro*) μ=0,9880 (σ=0,0216), dan F1 (*Macro*) μ=0,9837 (σ=0,0219). Pola distribusi Model E memiliki kemiripan yang sangat erat dengan Model A, yakni kotak IQR Akurasi yang sempit di rentang tinggi, kotak IQR Presisi (*Macro*) dan F1 (*Macro*) yang lebar dengan whisker panjang, serta keberadaan sebuah titik pencilan pada metrik *Recall* (*Macro*). Nilai σ yang identik dengan Model A pada seluruh metrik

mengonfirmasi bahwa kedua model memiliki tingkat kestabilan dan fluktuasi yang setara. Kondisi ini mengindikasikan bahwa Model E stabil pada Akurasi, namun masih mengalami fluktuasi yang perlu diperhatikan pada metrik Presisi dan F1 di antara *fold* pengujian.

4.1.6 Model F

Model F dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (128, 64, 32) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

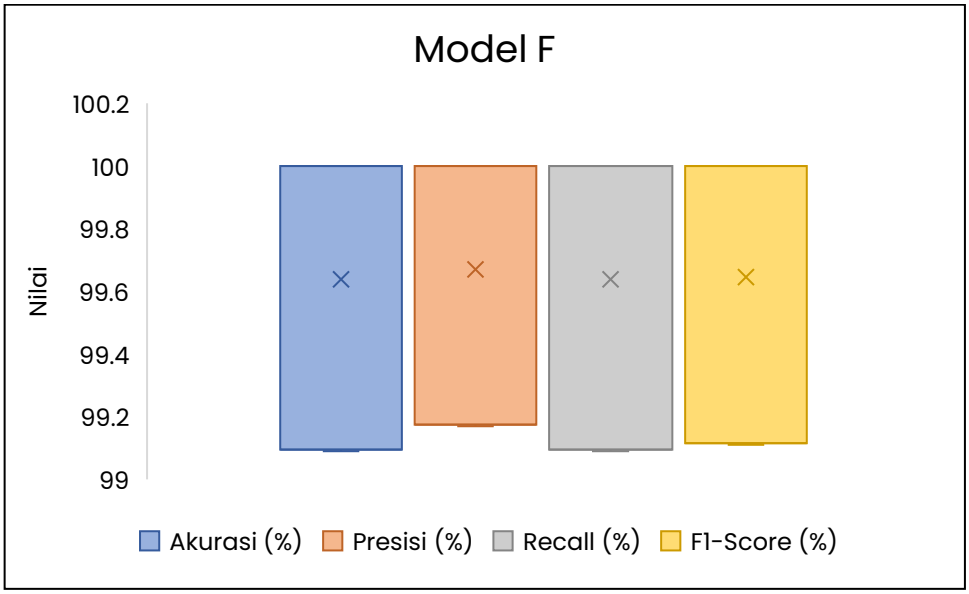


Gambar 4. 12 Metrik Evaluasi Model F

Tabel 4. 6 Metrik Evaluasi Model F

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.12, Model F mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 13 Visualisasi Rata-rata evaluation *matrix* Model F

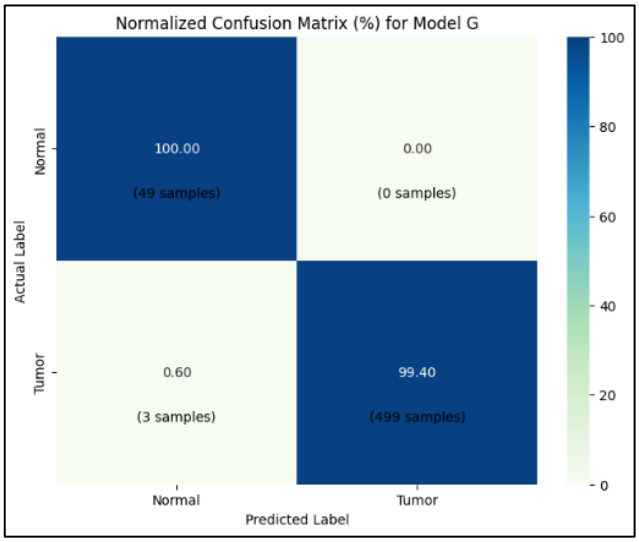
Lebih lanjut, distribusi performa Model F divisualisasikan melalui grafik *boxplot* pada Gambar 4.13. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*)

μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model F memperlihatkan pola yang identik dengan Model B, di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi, mencerminkan konsistensi yang sangat baik dengan σ *Recall* hanya sebesar 0,0024. Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah yang panjang, menunjukkan variasi yang substansial pada dimensi presisi di antara *fold* pengujian. Secara keseluruhan, Model F menunjukkan kestabilan Akurasi dan *Recall* yang sangat tinggi, namun konsistensi pada Presisi (*Macro*) masih menjadi tantangan yang melekat pada konfigurasi ini.

4.1.7 Model G

Model G dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (128, 64, 32) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

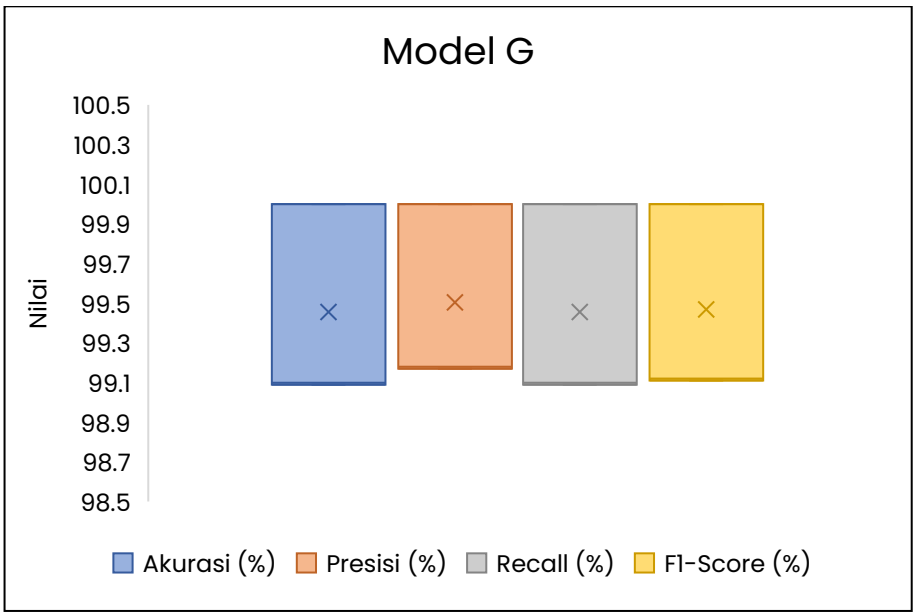


Gambar 4. 14 Metrik Evaluasi Model G

Tabel 4. 7 Metrik Evaluasi Model G

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**99,51**	**99,46**	**99,47**

Berdasarkan Gambar 4.14, Model G mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 498 data kanker sebagai *true positive*, dengan 3 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 99,49%, *recall* 99,46%, dan F1-score 99,46%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



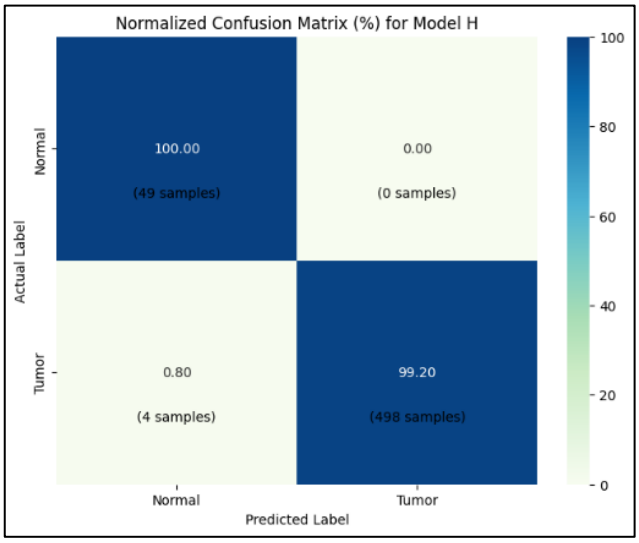
Gambar 4. 15 Visualisasi Rata-rata evaluation *matrix* Model G

Lebih lanjut, distribusi performa Model G divisualisasikan melalui grafik *boxplot* pada Gambar 4.15. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0045), Presisi (*Macro*) μ=0,9727 (σ=0,0223), *Recall* (*Macro*) μ=0,9970 (σ=0,0024), dan F1 (*Macro*) μ=0,9842 (σ=0,0129). Kotak IQR pada Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi, dengan σ *Recall* sebesar 0,0024 yang mencerminkan kestabilan yang sangat tinggi pada kemampuan deteksi model. Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar dengan whisker bawah yang panjang, mencerminkan adanya variasi yang cukup besar pada dimensi presisi antar*fold* , sebagaimana dikonfirmasi oleh nilai σ=0,0223. Secara keseluruhan, Model G sangat stabil pada Akurasi dan *Recall*, namun fluktuasi pada Presisi (*Macro*) mengindikasikan bahwa konsistensi pada metrik tersebut belum sepenuhnya terjaga.

4.1.8 Model H

Model H dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (128, 64, 32) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

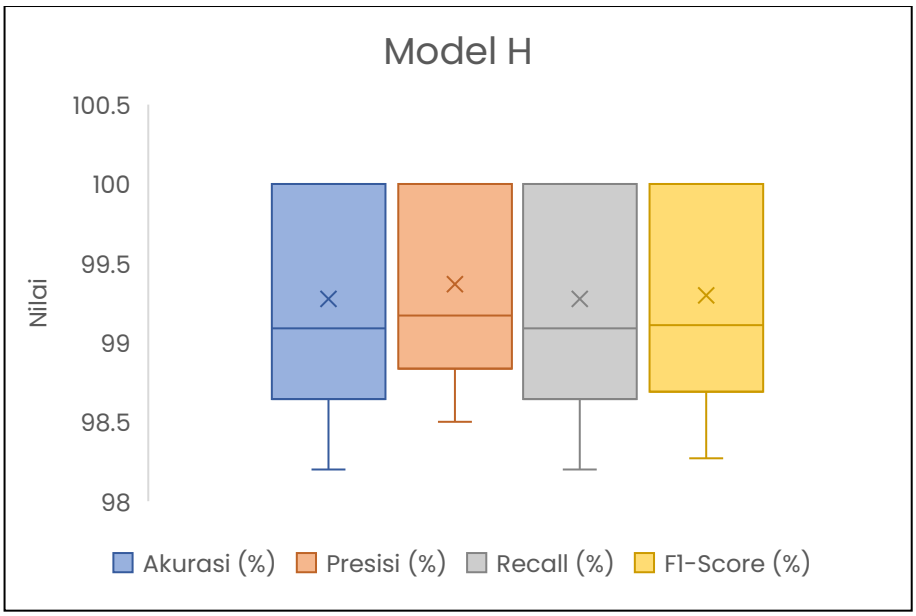


Gambar 4. 16 Metrik Evaluasi Model H

Tabel 4. 8 Metrik Evaluasi Model H

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,20	98,50	98,20	98,27
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,28**	**99,37**	**99,28**	**99,30**

Berdasarkan Gambar 4.16, Model H mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 497 data kanker sebagai *true positive*, dengan 4 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,28%, presisi 99,33%, *recall* 99,27%, dan F1-score 99,29%. Nilai standar deviasi sebesar 0,0068 menunjukkan konsistensi performa yang cukup stabil.



Gambar 4. 17 Visualisasi Rata-rata evaluation *matrix* Model H

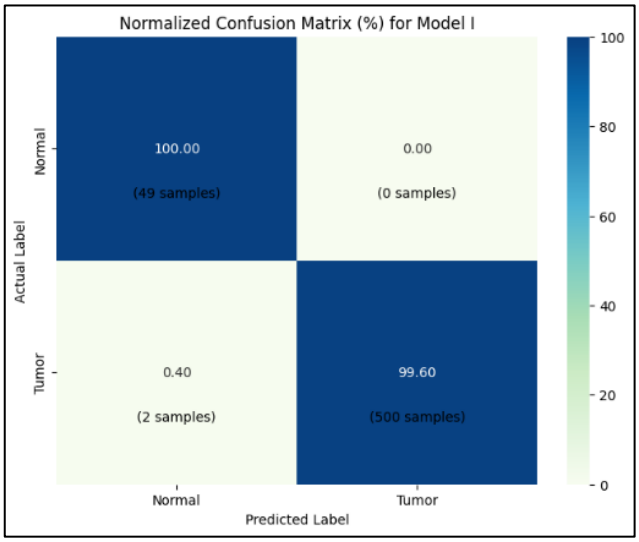
Lebih lanjut, distribusi performa Model H divisualisasikan melalui grafik *boxplot* pada Gambar 4.17. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9927 (σ=0,0068), Presisi (*Macro*) μ=0,9652 (σ=0,0316), *Recall* (*Macro*) μ=0,9960 (σ=0,0037), dan F1 (*Macro*) μ=0,9794 (σ=0,0190). Pola distribusi Model H memiliki kemiripan yang sangat erat dengan Model C, di mana kotak IQR *Recall* (*Macro*) tampak sangat tipis mencerminkan kestabilan tinggi (σ=0,0037), sementara kotak IQR Presisi (*Macro*) tampak sangat lebar dengan whisker bawah

yang panjang hingga hampir mencapai nilai 0,92, sesuai dengan nilai σ=0,0316 yang terbesar di antara seluruh metrik Model H. Kondisi ini mengindikasikan bahwa meskipun Model H menggunakan arsitektur yang lebih besar dengan 1.500 fitur, pola kestabilan dan fluktuasi yang dihasilkan pada dasarnya serupa dengan Model C, di mana *Recall* sangat stabil namun variasi Presisi masih menjadi catatan penting.

4.1.9 Model I

Model I dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (256, 128, 64) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

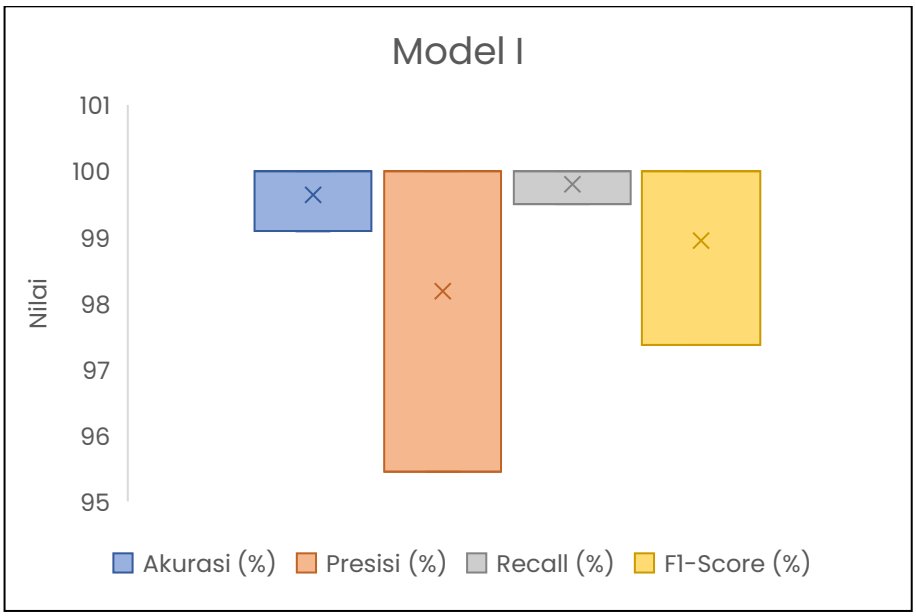


Gambar 4. 18 Metrik Evaluasi Model I

Tabel 4. 9 Metrik Evaluasi Model I

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99.10	95.45	99.50	97.37
Fold 2	100.00	100.00	100.00	100.00
Fold 3	99.09	95.45	99.50	97.37
Fold 4	100.00	100.00	100.00	100.00
Fold 5	100.00	100.00	100.00	100.00
Rata-rata	**99.64**	**98.18**	**99.80**	**98.95**

Berdasarkan Gambar 4.18, Model I mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



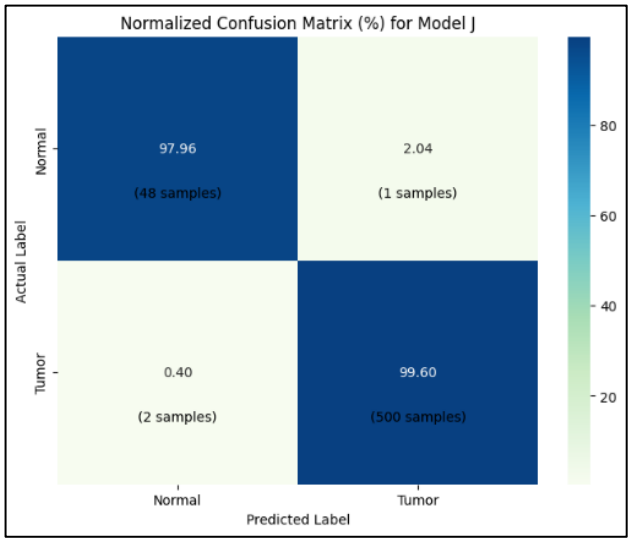
Gambar 4. 19 Visualisasi Rata-rata evaluation *matrix* Model I

Lebih lanjut, distribusi performa Model I divisualisasikan melalui grafik *boxplot* pada Gambar 4.19. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model I memperlihatkan pola yang identik dengan Model B dan F, yakni kotak IQR Akurasi dan *Recall* (*Macro*) yang sempit di rentang nilai tinggi dengan σ *Recall* sebesar 0,0024 yang sangat kecil, serta kotak IQR Presisi (*Macro*) yang sangat lebar disertai whisker bawah panjang. Kesamaan pola ini mengindikasikan bahwa meskipun Model I menggunakan arsitektur neuron yang lebih besar (256, 128, 64) dengan 50 fitur, karakteristik kestabilan dan fluktuasi yang dihasilkan tidak berbeda secara fundamental dari konfigurasi yang lebih kecil pada kelompok yang sama.

4.1.10 Model J

Model J dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (256, 128, 64) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

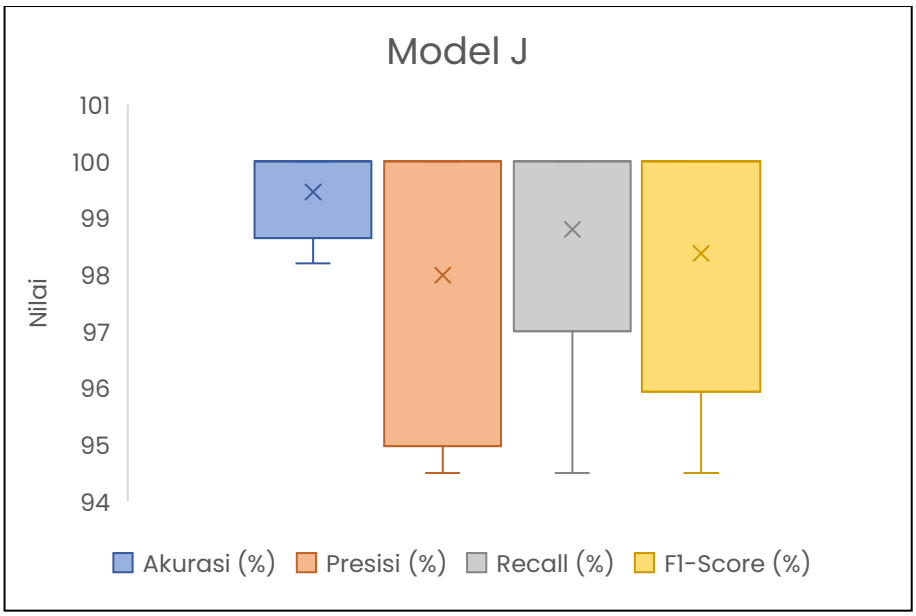


Gambar 4. 20 Metrik Evaluasi Model J

Tabel 4. 10 Metrik Evaluasi Model J

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,20	94,50	94,50	94,50
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	95,45	99,50	97,37
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**97,99**	**98,80**	**98,37**

Berdasarkan Gambar 4.20, Model J mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 499 data kanker sebagai *true positive*, dengan 1 *false negative* dan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 97,99%, *recall* 98,80%, dan F1-score 99,37%. Nilai standar deviasi sebesar 0,0072 menunjukkan adanya fluktuasi performa yang perlu diperhatikan di antara *fold* pengujian.



Gambar 4. 21 Visualisasi Rata-rata evaluation *matrix* Model J

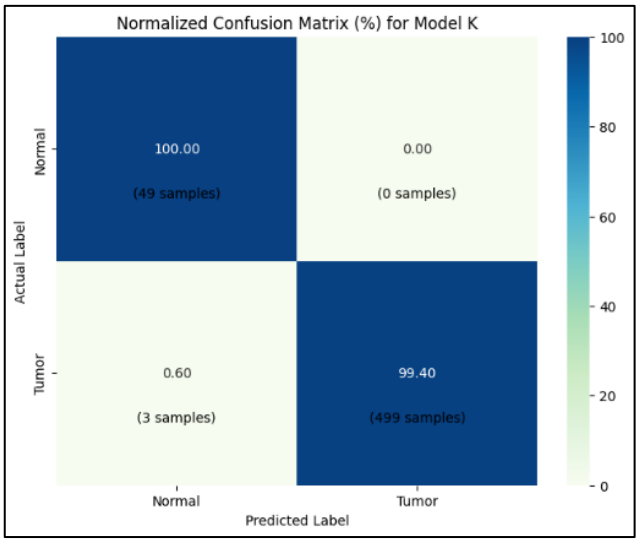
Lebih lanjut, distribusi performa Model J divisualisasikan melalui grafik *boxplot* pada Gambar 4.21. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0073), Presisi (*Macro*) μ=0,9799 (σ=0,0248), *Recall* (*Macro*) μ=0,9880 (σ=0,0216), dan F1 (*Macro*) μ=0,9837 (σ=0,0219). Distribusi performa Model J memperlihatkan pola yang identik dengan Model A dan E, yakni kotak IQR Akurasi yang sempit, kotak IQR Presisi (*Macro*) dan F1 (*Macro*) yang lebar dengan whisker panjang ke bawah, serta keberadaan sebuah titik pencilan pada metrik *Recall* (*Macro*). Kesamaan karakteristik distribusi antara ketiga model

tersebut mengindikasikan bahwa penggunaan 100 fitur pada arsitektur neuron (256, 128, 64) menghasilkan pola kestabilan yang setara dengan arsitektur-arsitektur lebih kecil pada jumlah fitur yang sama, sehingga disimpulkan bahwa kestabilan Akurasi terjaga namun fluktuasi pada Presisi dan F1 tetap menjadi catatan.

4.1.11 Model K

Model K dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (256, 128, 64) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information (*MI*)*.

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

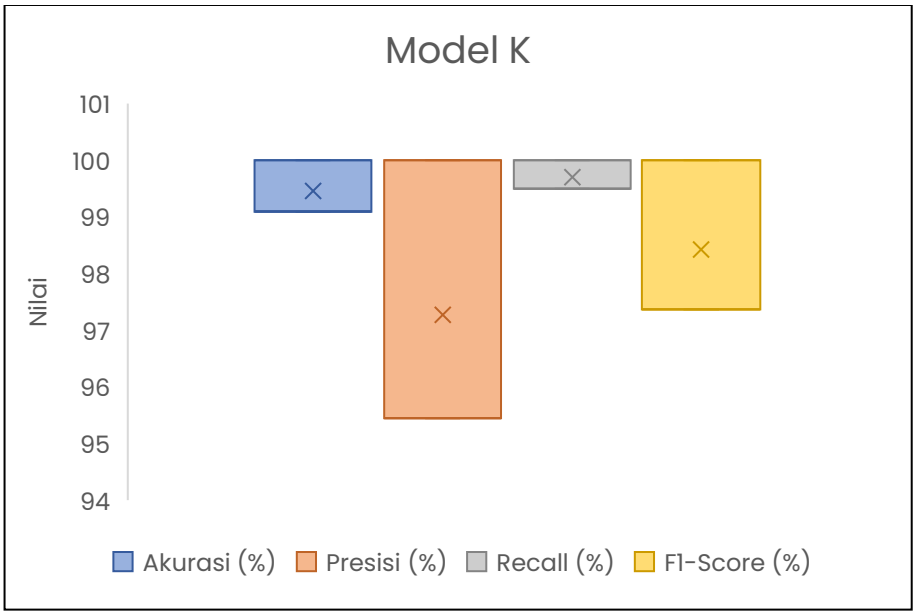


Gambar 4. 22 Metrik Evaluasi Model K

Tabel 4. 11 Metrik Evaluasi Model K

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	95,45	99,50	97,37
Fold 2	99,09	95,45	99,50	97,37
Fold 3	99,09	95,45	99,50	97,37
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**97,27**	**99,70**	**98,42**

Berdasarkan Gambar 4.22, Model K mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 498 data kanker sebagai *true positive*, dengan 3 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 99,49%, *recall* 99,46%, dan F1-score 99,46%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 23 Visualisasi Rata-rata evaluation *matrix* Model K

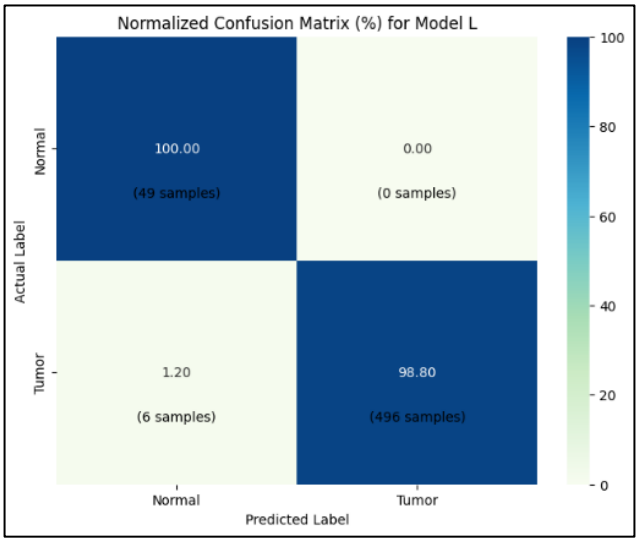
Lebih lanjut, distribusi performa Model K divisualisasikan melalui grafik *boxplot* pada Gambar 4.23. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0045), Presisi (*Macro*) μ=0,9727 (σ=0,0223), *Recall* (*Macro*)

μ=0,9970 (σ=0,0024), dan F1 (*Macro*) μ=0,9842 (σ=0,0129). Distribusi performa Model K memperlihatkan pola yang sangat serupa dengan Model G dan O, di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi dengan σ *Recall* yang sangat kecil sebesar 0,0024, mencerminkan kestabilan yang sangat tinggi pada kedua metrik tersebut. Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah yang panjang, mengindikasikan variasi yang substansial pada dimensi presisi dengan σ=0,0223. Kondisi ini mengonfirmasi bahwa peningkatan jumlah fitur hingga 500 pada arsitektur neuron (256, 128, 64) tidak mengubah pola distribusi performa secara fundamental dibandingkan konfigurasi-konfigurasi yang lebih kecil.

4.1.12 Model L

Model L dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (256, 128, 64) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

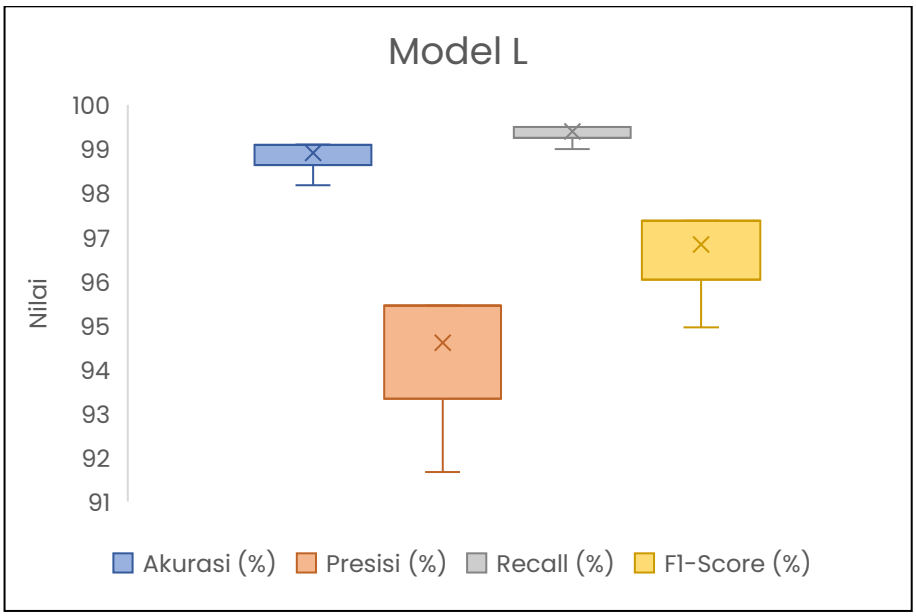


Gambar 4. 24 Metrik Evaluasi Model L

Tabel 4. 12 Metrik Evaluasi Model L

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	98,18	98,48	98,18	98,26
Fold 4	99,09	99,17	99,09	99,11
Fold 5	99,09	99,18	99,09	99,11
Rata-rata	**98,91**	**99,04**	**98,91**	**98,94**

Berdasarkan Gambar 4.24, Model L mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 495 data kanker sebagai *true positive*, dengan 6 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 98,91%, presisi 99,03%, *recall* 98,91%, dan F1-score 98,94%. Nilai standar deviasi sebesar 0,0036 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



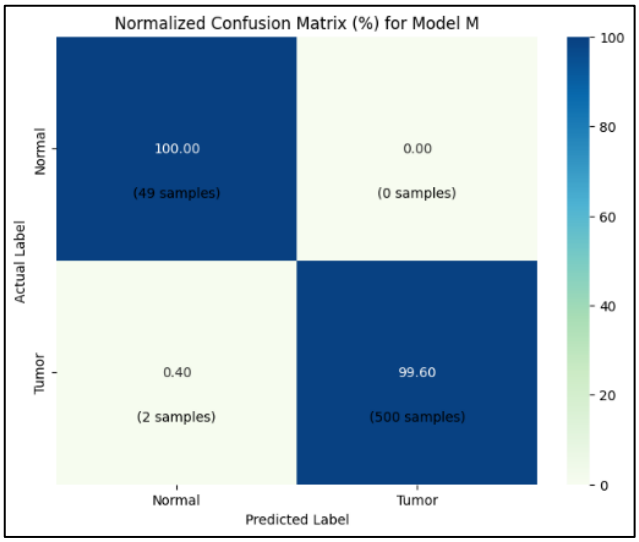
Gambar 4. 25 Visualisasi Rata-rata evaluation *matrix* Model L

Lebih lanjut, distribusi performa Model L divisualisasikan melalui grafik *boxplot* pada Gambar 4.25. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9891 (σ=0,0036), Presisi (*Macro*) μ=0,9461 (σ=0,0148), *Recall* (*Macro*) μ=0,9940 (σ=0,0020), dan F1 (*Macro*) μ=0,9683 (σ=0,0095). Grafik *boxplot* Model L memperlihatkan kotak IQR yang sangat sempit atau hampir tidak tampak pada seluruh metrik, namun disertai beberapa titik pencilan yang tersebar di luar batas whisker pada metrik Akurasi, *Recall* (*Macro*), dan F1 (*Macro*). Meskipun nilai σ pada sebagian besar metrik tergolong kecil, keberadaan titik-titik pencilan ini mengindikasikan bahwa performa model pada *fold -fold* tertentu mengalami penurunan yang signifikan dari nilai umumnya. Secara keseluruhan, Model L menunjukkan kestabilan yang baik pada sebagian besar *fold* , namun rentan terhadap fluktuasi pada kondisi partisi data tertentu, terutama pada metrik Presisi (*Macro*).

4.1.13 Model M

Model M dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

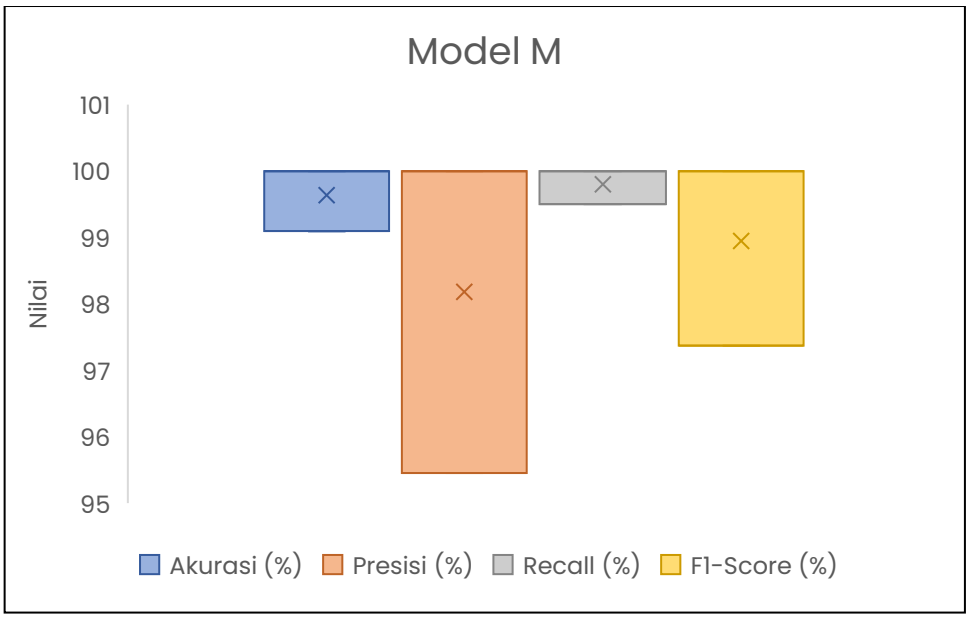


Gambar 4. 26 Metrik Evaluasi Model M

Tabel 4. 13 Metrik Evaluasi Model M

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.26, Model M mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,67%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 27 Visualisasi Rata-rata evaluation *matrix* Model M

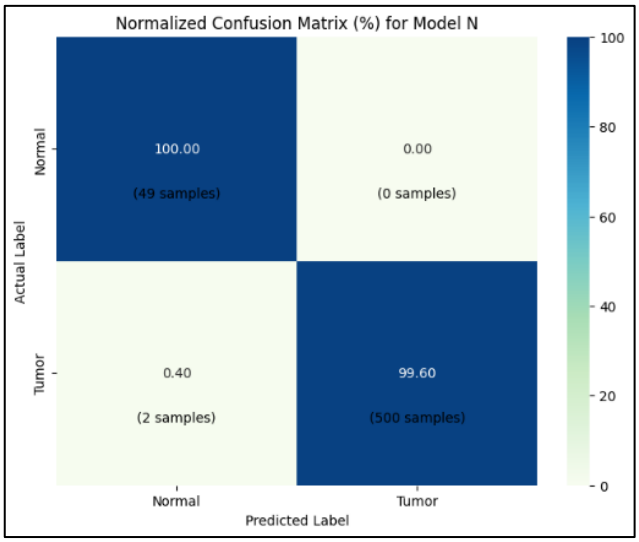
Lebih lanjut, distribusi performa Model M divisualisasikan melalui grafik *boxplot* pada Gambar 4.27. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model M memperlihatkan pola yang serupa dengan kelompok model berkonfigurasi dua *hidden layer*, yakni kotak IQR Akurasi dan *Recall* (*Macro*) yang sempit di rentang nilai tinggi mencerminkan kestabilan yang baik, sementara kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah yang panjang.

Nilai σ *Recall* (*Macro*) sebesar 0,0024 yang sangat kecil mengonfirmasi konsistensi tinggi pada kemampuan deteksi model. Secara keseluruhan, Model M menunjukkan kestabilan yang sangat baik pada Akurasi dan *Recall*, meskipun variasi pada Presisi (*Macro*) tetap menjadi karakteristik yang melekat pada konfigurasi ini.

4.1.14 Model N

Model N dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

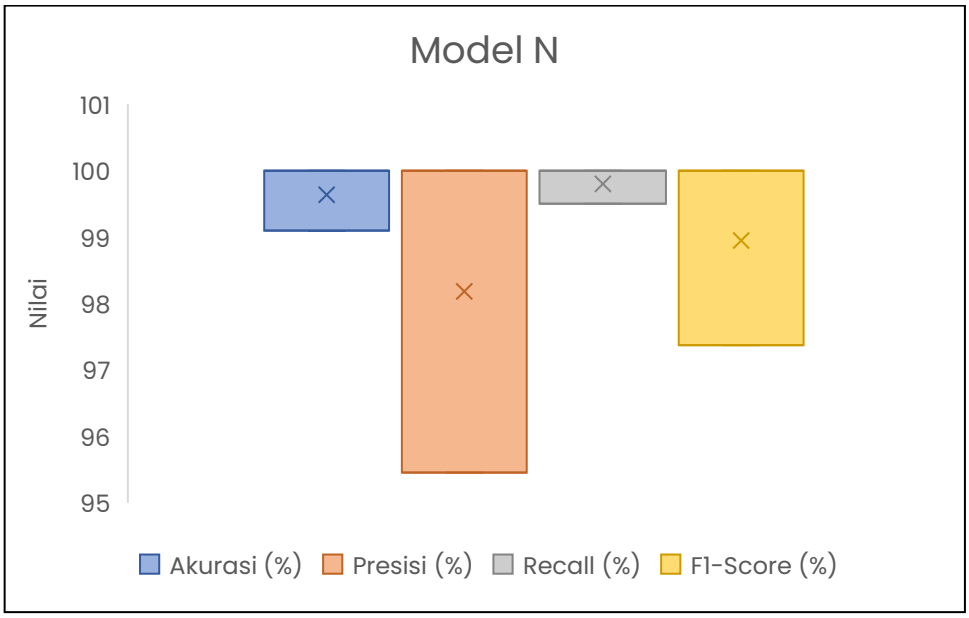


Gambar 4. 28 Metrik Evaluasi Model N

Tabel 4. 14 Metrik Evaluasi Model N

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.28, Model N mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 29 Visualisasi Rata-rata evaluation *matrix* Model N

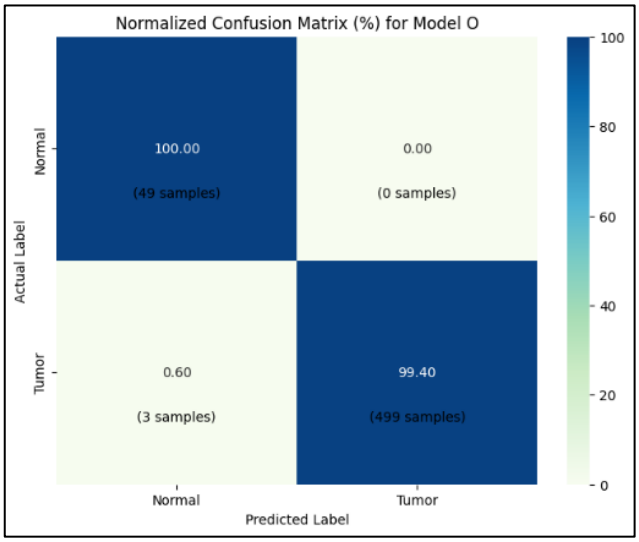
Lebih lanjut, distribusi performa Model N divisualisasikan melalui grafik *boxplot* pada Gambar 4.29. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*)

μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model N identik dengan Model M, di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi, sedangkan kotak IQR Presisi (*Macro*) tampak sangat lebar dengan whisker bawah yang panjang. Kesamaan pola distribusi antara Model M dan N mengindikasikan bahwa peningkatan jumlah fitur dari 50 menjadi 100 pada arsitektur dua *hidden layer* (32, 16) tidak memberikan perubahan yang berarti pada tingkat kestabilan dan pola fluktuasi performa di antara *fold* pengujian.

4.1.15 Model O

Model O dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

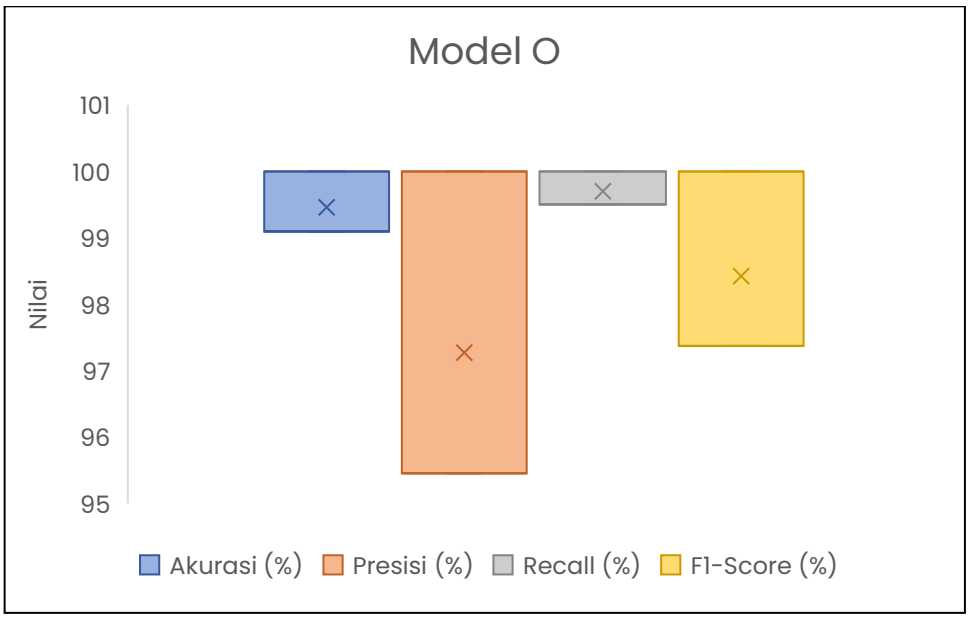


Gambar 4. 30 Metrik Evaluasi Model O

Tabel 4. 15 Metrik Evaluasi Model O

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**99,51**	**99,46**	**99,47**

Berdasarkan Gambar 4.30, Model O mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 498 data kanker sebagai *true positive*, dengan 3 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 99,49%, *recall* 99,46%, dan F1-score 99,46%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



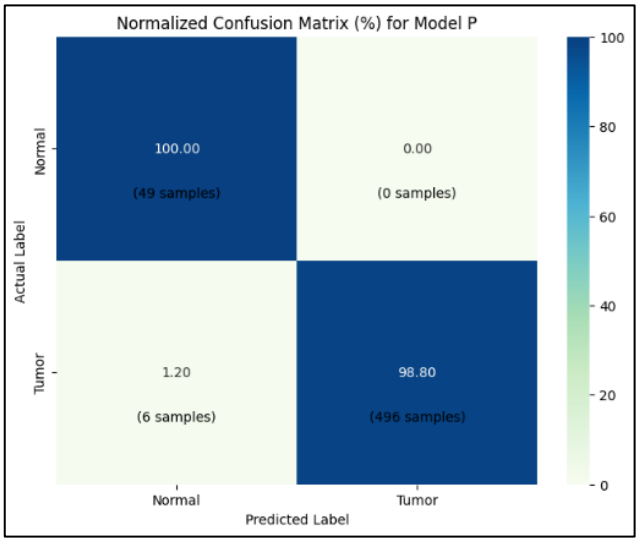
Gambar 4. 31 Visualisasi Rata-rata evaluation *matrix* Model O

Lebih lanjut, distribusi performa Model O divisualisasikan melalui grafik *boxplot* pada Gambar 4.31. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0045), Presisi (*Macro*) μ=0,9727 (σ=0,0223), *Recall* (*Macro*) μ=0,9970 (σ=0,0024), dan F1 (*Macro*) μ=0,9842 (σ=0,0129). Distribusi performa Model O memperlihatkan pola yang sangat serupa dengan Model G dan K, yakni kotak IQR Akurasi dan *Recall* (*Macro*) yang sempit di rentang nilai tinggi dengan σ *Recall* sebesar 0,0024, serta kotak IQR Presisi (*Macro*) yang sangat lebar disertai whisker bawah panjang. Konsistensi pola distribusi pada ketiga model ini mengindikasikan bahwa variasi jumlah fitur pada arsitektur yang sama tidak mengubah karakteristik kestabilan secara fundamental, di mana Akurasi dan *Recall* tetap sangat stabil namun variasi pada Presisi (*Macro*) tetap menjadi catatan yang perlu diperhatikan.

4.1.16 Model P

Model P dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

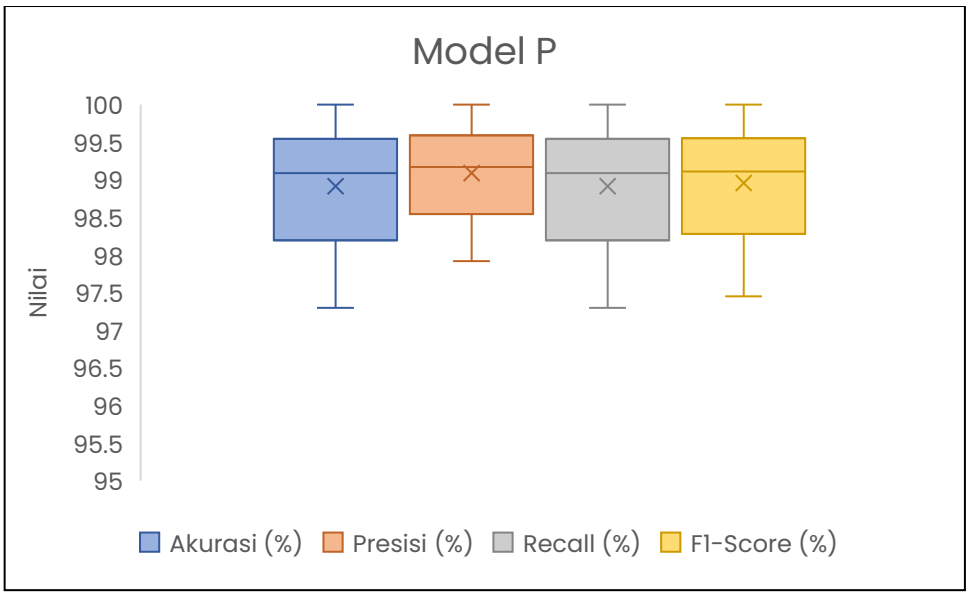


Gambar 4. 32 Metrik Evaluasi Model P

Tabel 4. 16 Metrik Evaluasi Model P

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	97.30	97.92	97.30	97.45
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	99,09	99,18	99,09	99,11
Rata-rata	**98,91**	**99,09**	**98,91**	**98,96**

Berdasarkan Gambar 4.32, Model P mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 495 data kanker sebagai *true positive*, dengan 6 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 98,91%, presisi 99,03%, *recall* 98,91%, dan F1-score 98,94%. Nilai standar deviasi sebesar 0,0088 menunjukkan adanya fluktuasi performa yang perlu diperhatikan di antara *fold* pengujian.



Gambar 4. 33 Visualisasi Rata-rata evaluation *matrix* Model P

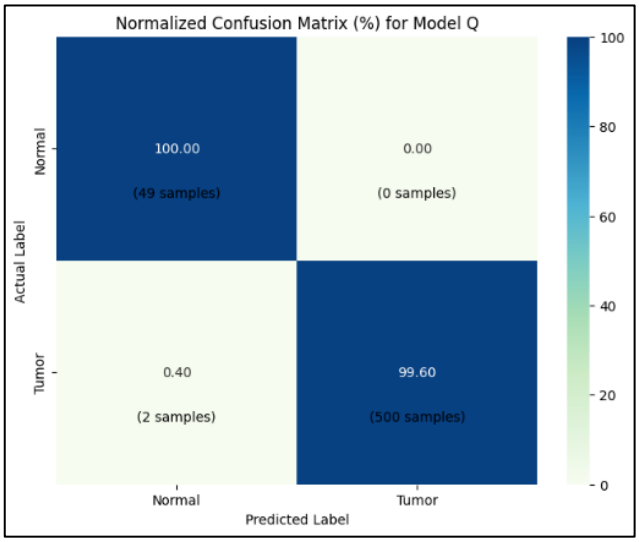
Lebih lanjut, distribusi performa Model P divisualisasikan melalui grafik *boxplot* pada Gambar 4.33. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9891 (σ=0,0089), Presisi (*Macro*) μ=0,9487 (σ=0,0369), *Recall* (*Macro*) μ=0,9940 (σ=0,0049), dan F1 (*Macro*) μ=0,9691 (σ=0,0235). Grafik *boxplot* Model P memperlihatkan kotak IQR yang sangat sempit pada seluruh metrik, namun hampir semua metrik disertai dengan beberapa titik pencilan yang tersebar di atas maupun di bawah batas *whisker*. Kondisi ini mengindikasikan bahwa sebagian besar *fold* menghasilkan performa yang seragam, namun pada *fold -fold* tertentu

model mengalami penurunan atau peningkatan performa yang cukup drastis. Nilai σ Presisi (*Macro*) sebesar 0,0369 yang terbesar di antara seluruh metrik mengonfirmasi bahwa fluktuasi paling signifikan terjadi pada dimensi presisi, sehingga Model P dikategorikan sebagai model dengan konsistensi yang tidak merata antar*fold* .

4.1.17 Model Q

Model Q dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (64, 32) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

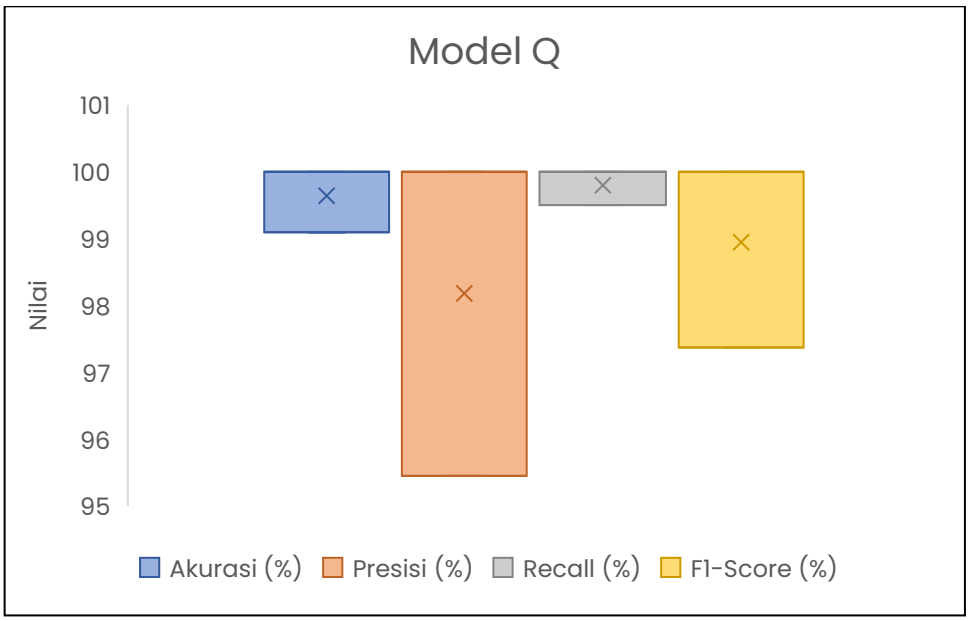


Gambar 4. 34 Metrik Evaluasi Model Q

Tabel 4. 17 Metrik Evaluasi Model Q

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**99,51**	**99,46**	**99,47**

Berdasarkan Gambar 4.34, Model Q mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



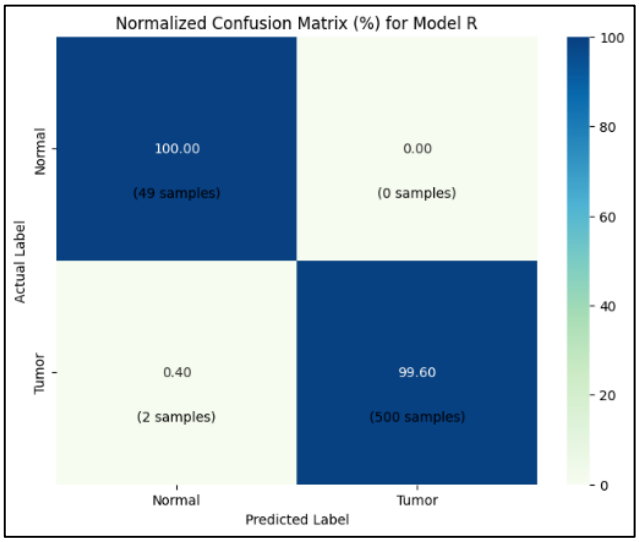
Gambar 4. 35 Visualisasi Rata-rata evaluation *matrix* Model Q

Lebih lanjut, distribusi performa Model Q divisualisasikan melalui grafik *boxplot* pada Gambar 4.35. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model Q memperlihatkan pola yang identik dengan sejumlah model berkonfigurasi dua *hidden layer* lainnya, yakni kotak IQR Akurasi dan *Recall* (*Macro*) yang sempit di rentang nilai tinggi, serta kotak IQR Presisi (*Macro*) yang sangat lebar dengan whisker bawah panjang. Nilai σ *Recall* sebesar 0,0024 mengonfirmasi konsistensi yang sangat tinggi pada kemampuan deteksi model. Secara keseluruhan, Model Q menunjukkan kestabilan yang sangat baik pada Akurasi dan *Recall*, sementara fluktuasi pada Presisi (*Macro*) dengan σ=0,0223 tetap menjadi karakteristik yang melekat pada konfigurasi ini.

4.1.18 Model R

Model R dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (64, 32) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

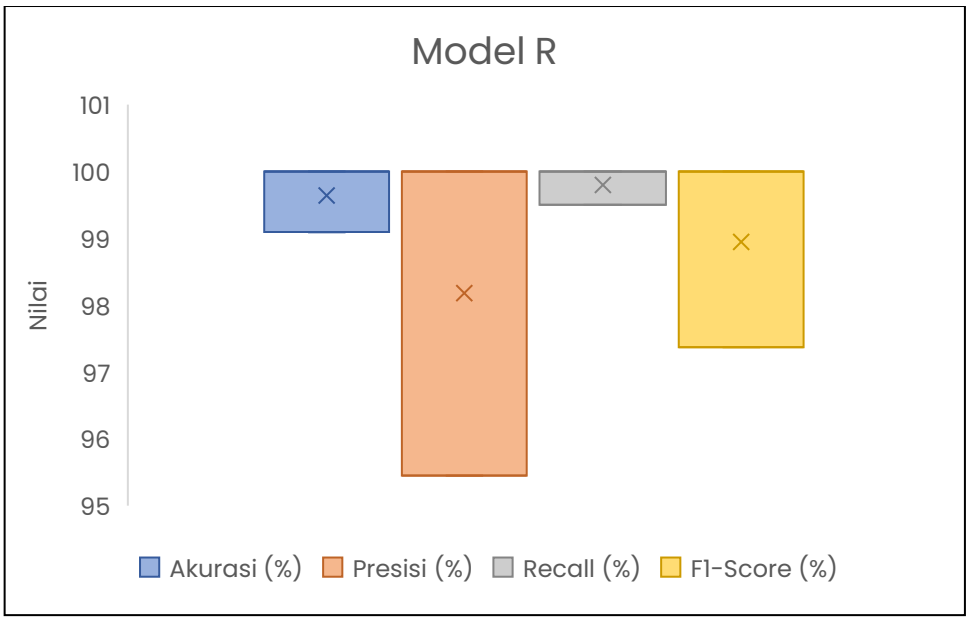


Gambar 4. 36 Metrik Evaluasi Model R

Tabel 4. 18 Metrik Evaluasi Model R

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.36, Model R mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 37 Visualisasi Rata-rata evaluation *matrix* Model R

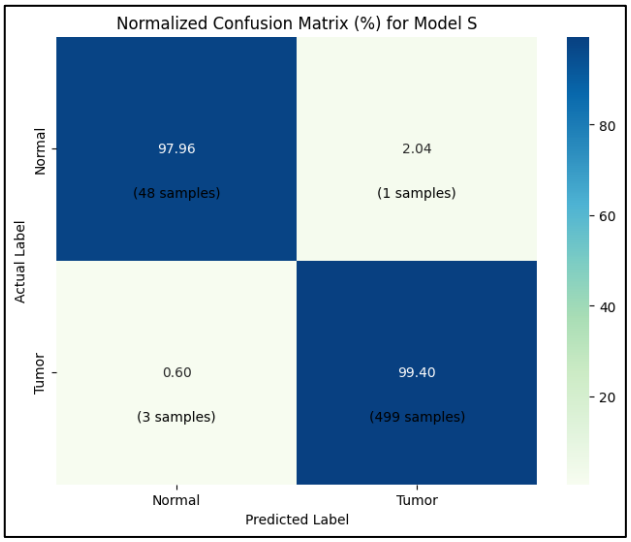
Lebih lanjut, distribusi performa Model R divisualisasikan melalui grafik *boxplot* pada Gambar 4.37. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model R identik dengan sejumlah model dalam kelompok yang sama, di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi mencerminkan konsistensi yang baik, sedangkan kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah panjang. Kesamaan pola ini

mengindikasikan bahwa penambahan jumlah fitur dari 50 menjadi 100 pada arsitektur dua *hidden layer* (64, 32) tidak mengubah pola distribusi performa secara fundamental, sehingga Model R menghasilkan tingkat kestabilan yang setara dengan konfigurasi-konfigurasi yang berdekatan.

4.1.19 Model S

Model S dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (64, 32) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

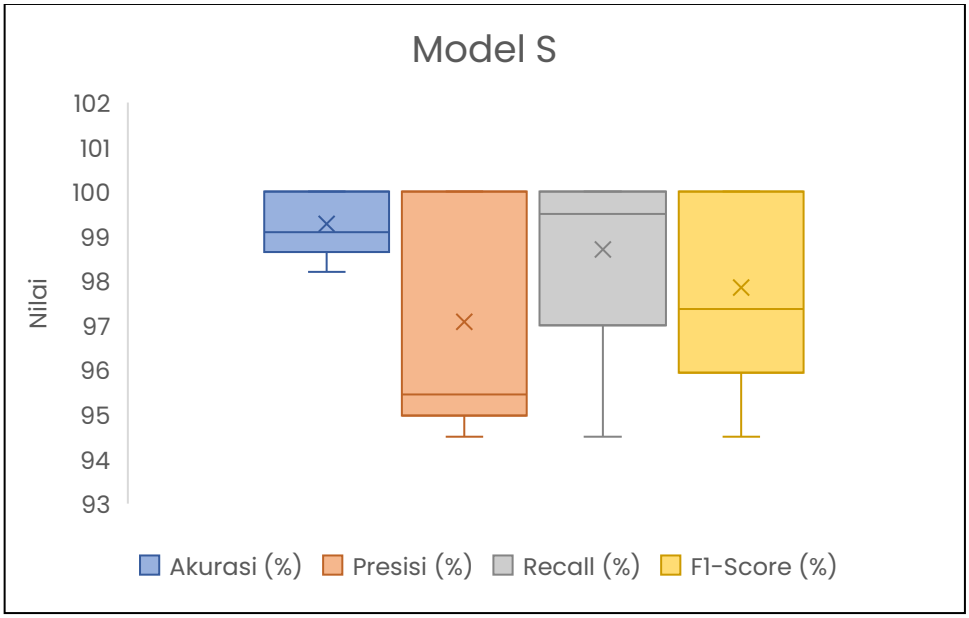


Gambar 4. 38 Metrik Evaluasi Model S

Tabel 4. 19 Metrik Evaluasi Model S

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,20	98,20	98,20	98,20
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,28**	**99,31**	**99,28**	**99,28**

Berdasarkan Gambar 4.38, Model S mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 498 data kanker sebagai *true positive*, dengan 1 *false negative* dan 3 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,28%, presisi 99,31%, *recall* 99,28%, dan F1-score 99,28%. Nilai standar deviasi sebesar 0,0068 menunjukkan konsistensi performa yang cukup stabil.



Gambar 4. 39 Visualisasi Rata-rata evaluation *matrix* Model S

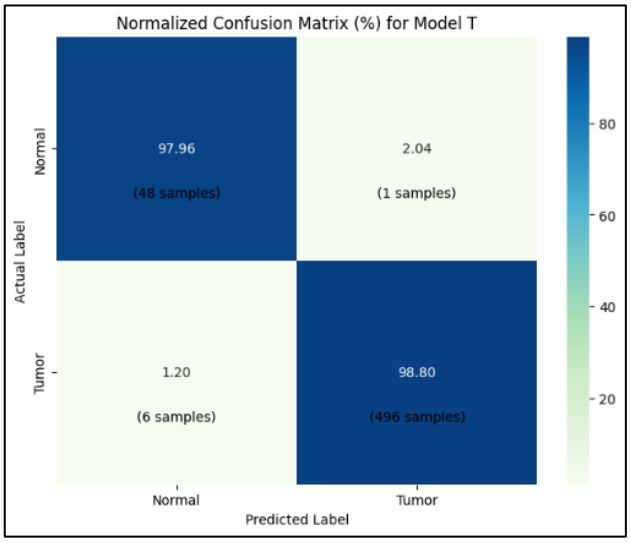
Lebih lanjut, distribusi performa Model S divisualisasikan melalui grafik *boxplot* pada Gambar 4.39. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9927 (σ=0,0068), Presisi (*Macro*) μ=0,9708 (σ=0,0241), *Recall* (*Macro*) μ=0,9870 (σ=0,0211), dan F1 (*Macro*) μ=0,9785 (σ=0,0205). Secara visual, kotak

IQR pada seluruh metrik tampak cukup lebar dengan whisker yang memanjang ke bawah, dan pada metrik *Recall* (*Macro*) terdapat sebuah titik pencilan di bawah batas whisker yang mengindikasikan adanya satu *fold* dengan performa yang sangat rendah. Berbeda dari sebagian besar model lainnya yang fluktuasinya terkonsentrasi pada Presisi (*Macro*), Model S memperlihatkan fluktuasi yang merata di hampir semua metrik sekaligus, sebagaimana tercermin dari nilai σ yang relatif serupa pada Presisi (0,0241), *Recall* (0,0211), dan F1 (0,0205). Kondisi ini mengindikasikan bahwa Model S memiliki tingkat konsistensi yang lebih rendah dibandingkan sebagian besar model lainnya.

4.1.20 Model T

Model T dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (64, 32) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

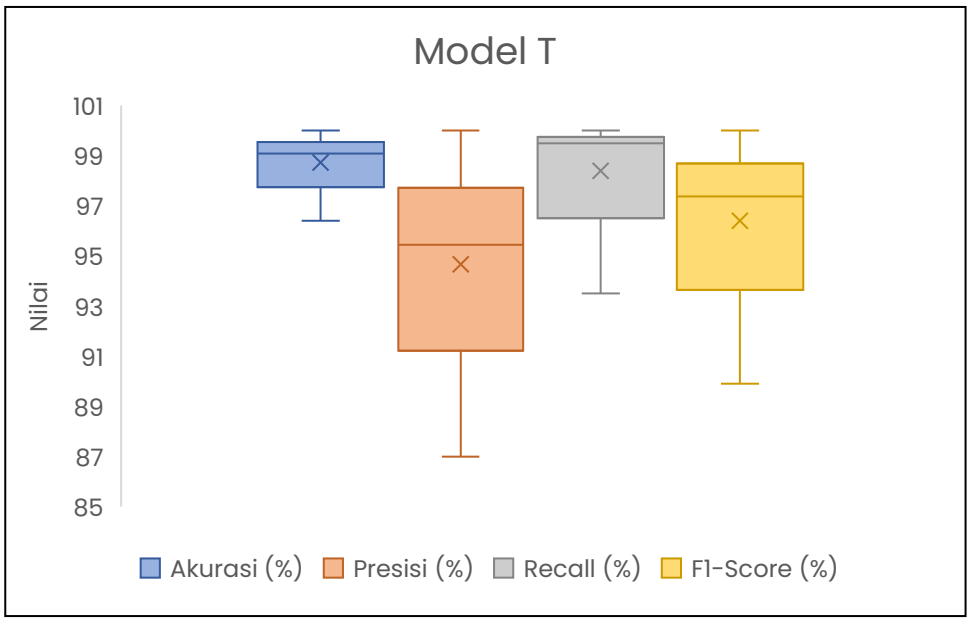


Gambar 4. 40 Metrik Evaluasi Model T

Tabel 4. 20 Metrik Evaluasi Model T

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	96.40	96.83	96.40	96.54
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	99,09	99,17	99,09	99,11
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**98,73**	**98,87**	**98,73**	**98,77**

Berdasarkan Gambar 4.40, Model T mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 495 data kanker sebagai *true positive*, dengan 1 *false negative* dan 6 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 98,73%, presisi 98,83%, *recall* 98,73%, dan F1-score 98,76%. Nilai standar deviasi sebesar 0,0122 merupakan yang tertinggi di antara seluruh model, mengindikasikan fluktuasi performa yang signifikan antar *fold* pengujian.



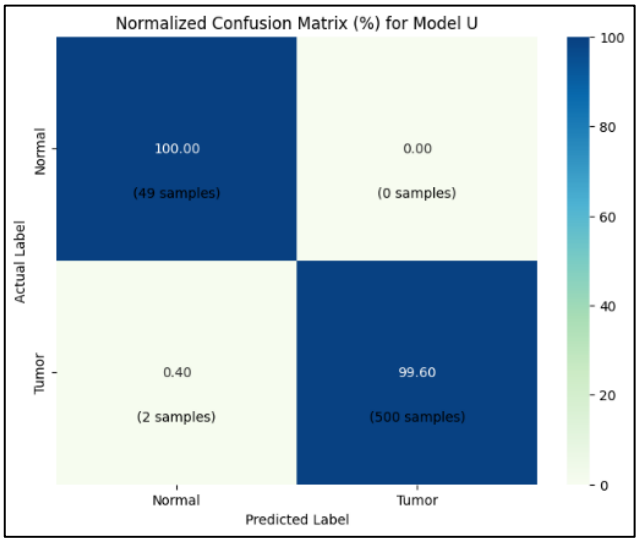
Gambar 4. 41 Visualisasi Rata-rata evaluation matrix *Model T*

Lebih lanjut, distribusi performa Model T divisualisasikan melalui grafik *boxplot* pada Gambar 4.41. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9873 (σ=0,0123), Presisi (*Macro*) μ=0,9467 (σ=0,0422), *Recall* (*Macro*) μ=0,9840 (σ=0,0246), dan F1 (*Macro*) μ=0,9640 (σ=0,0341). Grafik *boxplot* Model T memperlihatkan kotak IQR yang hampir tidak tampak pada seluruh metrik, namun hampir semua metrik disertai sejumlah titik pencilan yang tersebar jauh di atas maupun di bawah batas whisker, mengindikasikan adanya *fold -fold* dengan performa yang sangat menyimpang dari nilai umumnya. Nilai σ Presisi (*Macro*) sebesar 0,0422 merupakan salah satu yang tertinggi di antara seluruh model, dan dikombinasikan dengan σ F1 (*Macro*) sebesar 0,0341, mengindikasikan bahwa Model T mengalami fluktuasi performa yang paling parah di antara seluruh skenario yang diuji sehingga keandalannya perlu dipertimbangkan secara hati-hati.

4.1.21 Model U

Model U dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (256, 128) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

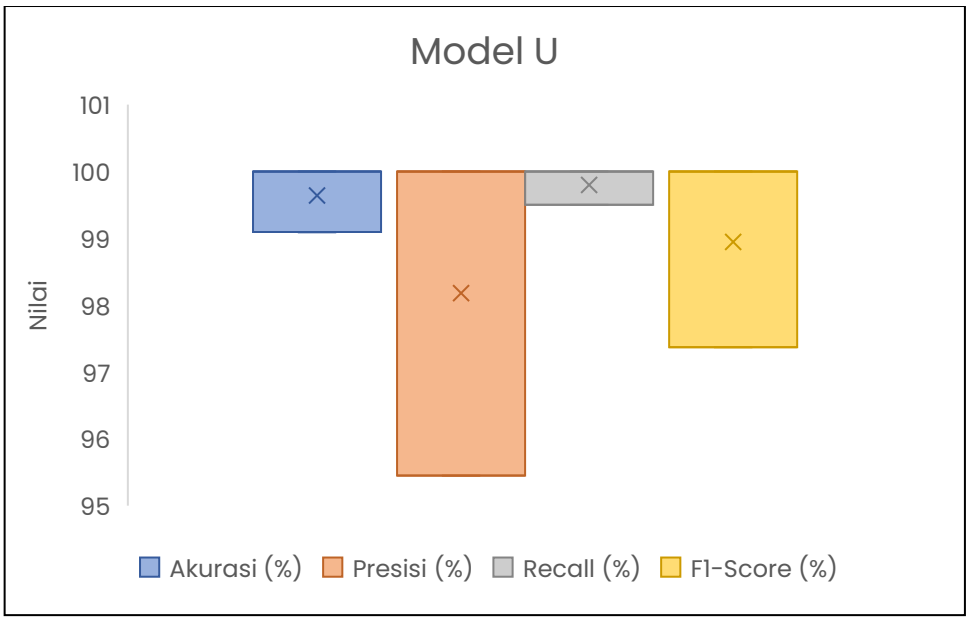


Gambar 4. 42 Metrik Evaluasi Model U

Tabel 4. 21 Metrik Evaluasi Model U

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.42, Model U mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 43 Visualisasi Rata-rata evaluation *matrix* Model U

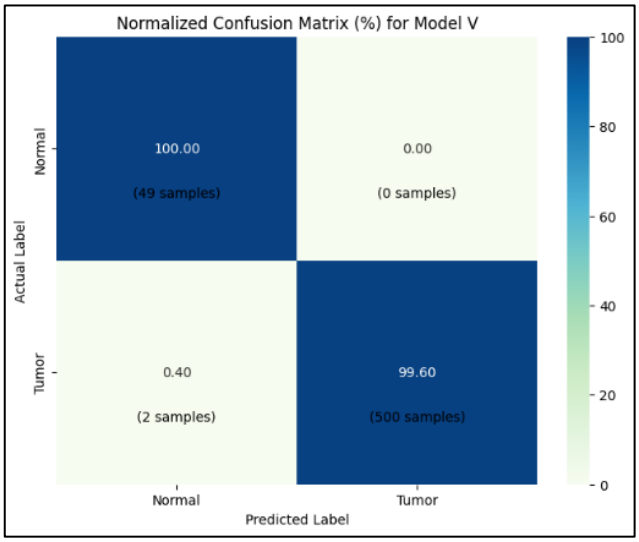
Lebih lanjut, distribusi performa Model U divisualisasikan melalui grafik *boxplot* pada Gambar 4.43. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model U memperlihatkan pola yang identik dengan sejumlah model berkonfigurasi dua *hidden layer* lainnya, yakni kotak IQR Akurasi dan *Recall* (*Macro*) yang sempit di rentang nilai tinggi dengan σ *Recall* yang sangat kecil sebesar 0,0024, mencerminkan kestabilan yang sangat tinggi pada kedua metrik tersebut.

Sebaliknya, kotak IQR Presisi (*Macro*) tampak sangat lebar disertai whisker bawah yang panjang hingga hampir mencapai nilai 0,95, mengindikasikan variasi yang substansial pada dimensi presisi antar*fold* , sebagaimana dikonfirmasi oleh nilai σ=0,0223. Secara keseluruhan, Model U menunjukkan kestabilan yang sangat baik pada Akurasi dan *Recall*, sementara fluktuasi pada Presisi (*Macro*) tetap menjadi karakteristik yang melekat pada konfigurasi arsitektur neuron (256, 128) dengan 50 fitur ini.

4.1.22 Model V

Model V dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (256, 128) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

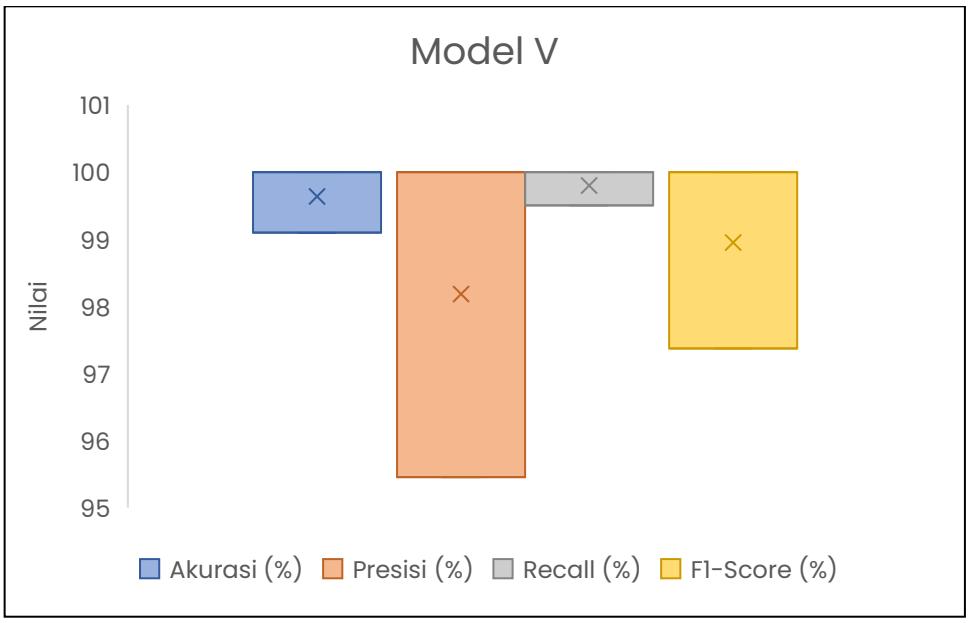


Gambar 4. 44 Metrik Evaluasi Model V

Tabel 4. 22 Metrik Evaluasi Model V

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	100,00	100,00	100,00	100,00
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,64**	**99,67**	**99,64**	**99,65**

Berdasarkan Gambar 4.44, Model V mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 499 data kanker sebagai *true positive*, dengan 2 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,64%, presisi 99,65%, *recall* 99,64%, dan F1-score 99,64%. Nilai standar deviasi sebesar 0,0044 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



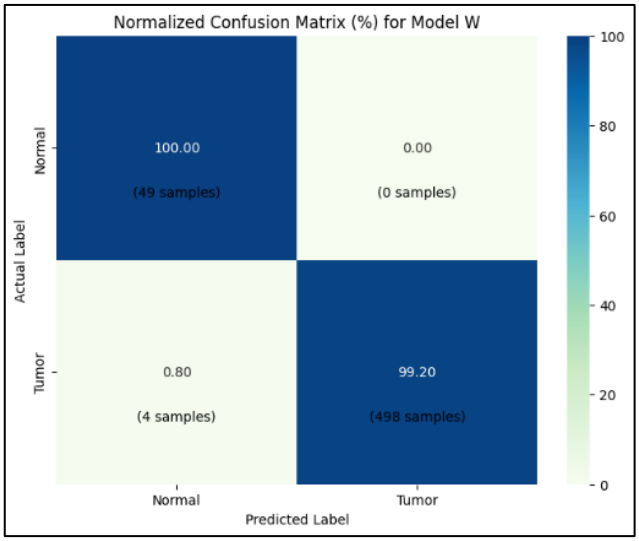
Gambar 4. 45 Visualisasi Rata-rata evaluation *matrix* Model V

Lebih lanjut, distribusi performa Model V divisualisasikan melalui grafik *boxplot* pada Gambar 4.45. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9964 (σ=0,0045), Presisi (*Macro*) μ=0,9818 (σ=0,0223), *Recall* (*Macro*) μ=0,9980 (σ=0,0024), dan F1 (*Macro*) μ=0,9895 (σ=0,0129). Distribusi performa Model V identik dengan sejumlah model dalam kelompok yang sama, di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit di rentang nilai tinggi dengan σ *Recall* yang sangat kecil sebesar 0,0024, sedangkan kotak IQR Presisi (*Macro*) tampak sangat lebar dengan whisker bawah panjang. Kesamaan distribusi ini mengindikasikan bahwa peningkatan ukuran arsitektur pada dua *hidden layer* (256, 128) dengan 100 fitur tidak menghasilkan perubahan pola kestabilan yang berbeda dari konfigurasi-konfigurasi yang lebih kecil, sehingga Model V menunjukkan kestabilan Akurasi dan *Recall* yang sangat tinggi dengan fluktuasi Presisi yang tetap menjadi karakteristik kelompok ini.

4.1.23 Model W

Model W dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (256, 128) dan jumlah fitur input sebesar 500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

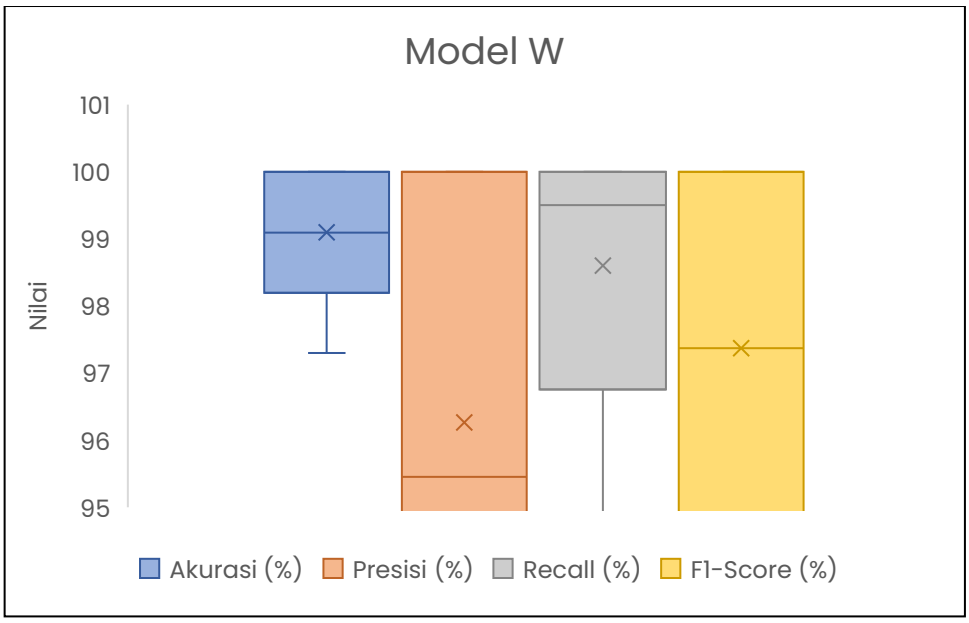


Gambar 4. 46 Metrik Evaluasi Model W

Tabel 4. 23 Metrik Evaluasi Model W

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,10	99,18	99,10	99,12
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	99,09	99,18	99,09	99,11
Rata-rata	**99,27**	**99,34**	**99,27**	**99,29**

Berdasarkan Gambar 4.46, Model W mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* (tanpa *false negative*) dan 497 data kanker sebagai *true positive*, dengan 4 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,27%, presisi 99,33%, *recall* 99,27%, dan F1-score 99,29%. Nilai standar deviasi sebesar 0,0036 menunjukkan konsistensi performa yang sangat stabil di antara lima *fold* pengujian.



Gambar 4. 47 Visualisasi Rata-rata evaluation *matrix* Model W

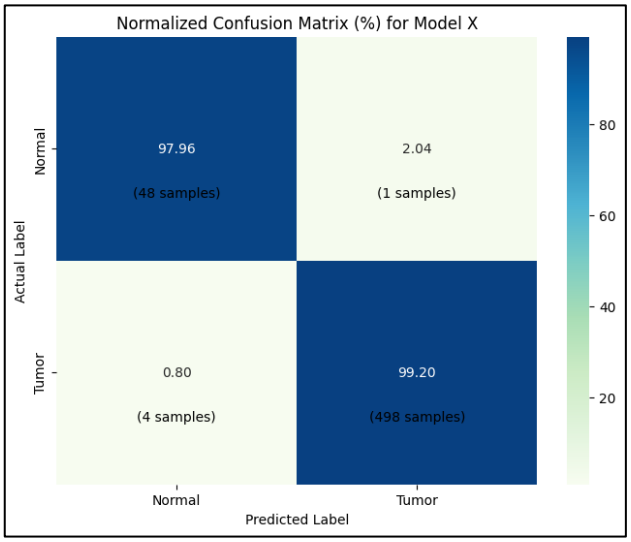
Lebih lanjut, distribusi performa Model W divisualisasikan melalui grafik *boxplot* pada Gambar 4.47. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9927 (σ=0,0036), Presisi (*Macro*) μ=0,9627 (σ=0,0187), *Recall* (*Macro*) μ=0,9960 (σ=0,0020), dan F1 (*Macro*) μ=0,9785 (σ=0,0108). Grafik *boxplot* Model W memperlihatkan kotak IQR yang sangat sempit atau hampir tidak tampak pada seluruh metrik, disertai beberapa titik pencilan yang tersebar di luar batas whisker pada metrik Presisi (*Macro*) dan F1 (*Macro*). Nilai σ Akurasi sebesar 0,0036 dan σ *Recall* (*Macro*) sebesar 0,0020 yang sangat kecil menunjukkan bahwa sebagian

besar *fold* menghasilkan performa yang sangat seragam pada kedua metrik tersebut, namun keberadaan titik pencilan mengindikasikan bahwa pada *fold* tertentu model masih mengalami penurunan performa. Secara keseluruhan, Model W sangat stabil pada Akurasi dan *Recall*, namun masih rentan terhadap fluktuasi pada dimensi Presisi dan F1 pada kondisi partisi data tertentu.

4.1.24 Model X

Model X dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (256, 128) dan jumlah fitur input sebesar 1500 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI).

Model ini memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 untuk mencegah *overfitting*, dikombinasikan dengan mekanisme *Early Stopping* yang menghentikan pelatihan secara otomatis ketika validasi *loss* tidak lagi membaik. Konfigurasi ini merupakan bagian dari skenario pengujian sistematis untuk menganalisis pengaruh kedalaman arsitektur dan jumlah fitur terhadap performa klasifikasi kanker paru LUSC.

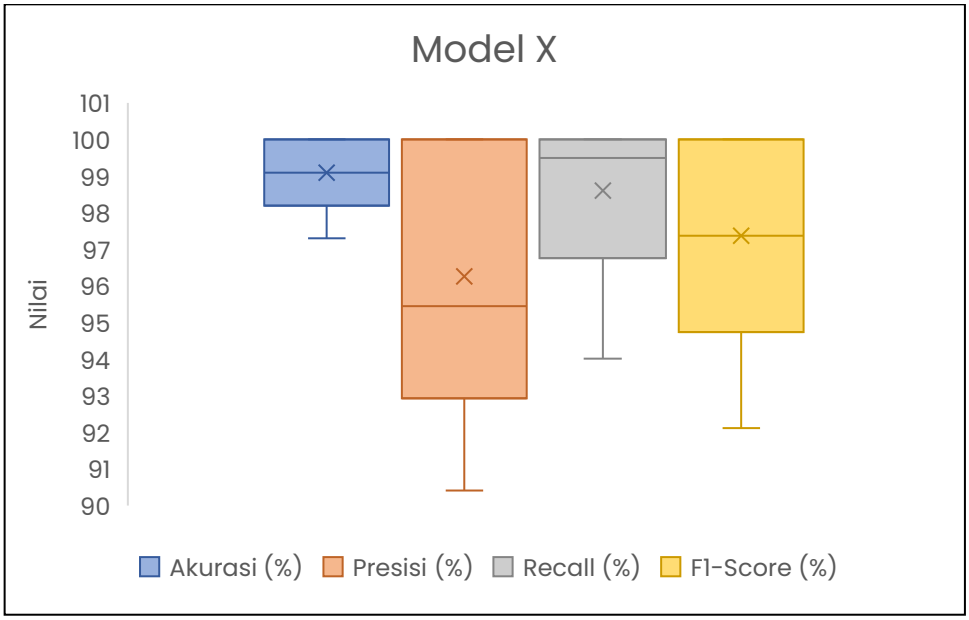


Gambar 4. 48 Metrik Evaluasi Model X

Tabel 4. 24 Metrik Evaluasi Model X

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	97.30	97.45	97.30	97.35
Fold 2	99,09	99,17	99,09	99,11
Fold 3	99,09	99,17	99,09	99,11
Fold 4	100,00	100,00	100,00	100,00
Fold 5	100,00	100,00	100,00	100,00
Rata-rata	**99,10**	**99,16**	**99,10**	**99,12**

Berdasarkan Gambar 4.48, Model X mampu mendeteksi 48 data jaringan normal sebagai *true negative* dan 497 data kanker sebagai *true positive*, dengan 1 *false negative* dan 4 *false positive*. Secara keseluruhan, model mencapai rata-rata akurasi 99,10%, presisi 99,13%, *recall* 99,09%, dan F1-score 99,10%. Nilai standar deviasi sebesar 0,0099 menunjukkan adanya fluktuasi performa yang perlu diperhatikan di antara *fold* pengujian.



Gambar 4. 49 Visualisasi Rata-rata evaluation *matrix* Model X

Lebih lanjut, distribusi performa Model X divisualisasikan melalui grafik *boxplot* pada Gambar 4.49. Berdasarkan grafik tersebut, diperoleh nilai Akurasi

μ=0,9909 (σ=0,0100), Presisi (*Macro*) μ=0,9626 (σ=0,0357), *Recall* (*Macro*) μ=0,9860 (σ=0,0231), dan F1 (*Macro*) μ=0,9737 (σ=0,0288). Secara visual, kotak IQR pada metrik Presisi (*Macro*) tampak sangat lebar dengan whisker yang menjangkau nilai sangat rendah, sesuai dengan nilai σ=0,0357 yang terbesar di antara seluruh metrik. Selain itu, terdapat titik-titik pencilan pada metrik Akurasi, *Recall* (*Macro*), dan F1 (*Macro*) yang menunjukkan adanya *fold -fold* dengan performa yang menyimpang secara signifikan. Nilai σ yang tinggi pada hampir seluruh metrik mengindikasikan bahwa Model X merupakan salah satu model dengan variabilitas performa tertinggi, sehingga arsitektur ini dinilai kurang stabil dalam menghadapi variasi partisi data yang berbeda-beda.

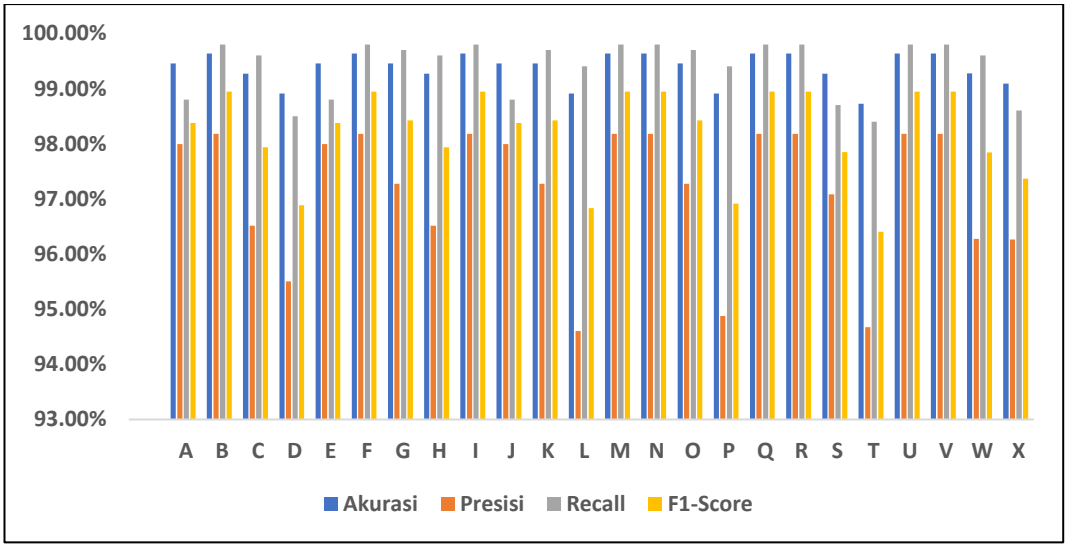
4.2 Pembahasan

Hasil pengujian dua puluh model MLP dengan variasi arsitektur jaringan, jumlah neuron, dan jumlah fitur dirangkum pada Tabel 4.21. Tabel tersebut menampilkan nilai rata-rata akurasi, presisi, *recall*, dan F1-score dari seluruh skenario yang telah diuji menggunakan metode 5-*Fold Cross validation*.

Hasil pengujian dua puluh empat model MLP dengan variasi arsitektur jaringan, jumlah neuron, dan jumlah fitur dirangkum pada Tabel 4.25. Tabel tersebut menampilkan nilai rata-rata akurasi, standar deviasi, presisi, *recall*, dan F1-score dari seluruh skenario yang telah diuji menggunakan metode 5-*Fold Cross validation*. Rentang akurasi yang diperoleh berada antara 98,73% hingga 99,64%, yang mengindikasikan bahwa kombinasi seleksi fitur berbasis *Mutual Information* dan arsitektur MLP mampu mempelajari pola ekspresi gen kanker paru sel skuamosa (LUSC) secara efektif pada semua konfigurasi yang diuji.

Tabel 4. 25 Rekapitulasi Hasil Pengujian Seluruh Model MLP

No	Model	*Hidden Layer*	Akurasi (*Mean*)	Presisi (*Mean*)	*Recall* (*Mean*)	F1-Score (*Mean*)
1	Model A	3	99,4545	97.9909	98,8	98,3736
2	Model B		99,6364	98,1818	99,8	98,9471
3	Model C		99,2727	96.5152	99,6	97.937
4	Model D		98,9091	95.5051	98,5	96.8841
5	Model E		99,4545	97.9909	98,8	98,3736
6	Model F		99,6364	98,1818	99,8	98,9471
7	Model G		99,4545	97.2727	99,7	98,4207
8	Model H		99,2727	96.5152	99,6	97.937
9	Model I		99,6364	98,1818	99,8	98,9471
10	Model J		99,4545	97.9909	98,8	98,3736
11	Model K		99,4545	97.2727	99,7	98,4207
12	Model L		98,9107	94.6061	99,402	96.835
13	Model M	2	99,6364	98,1818	99,8	98,9471
14	Model N		99,6364	98,1818	99,8	98,9471
15	Model O		99,4545	97.2727	99,7	98,4207
16	Model P		98,9107	94.8741	99,402	96.9149
17	Model Q		99,6364	98,1818	99,8	98,9471
18	Model R		99,6364	98,1818	99,8	98,9471
19	Model S		99,2727	97.0818	98,7	97.8471
20	Model T		98,7273	94.6707	98,4	96.4005
21	Model U		99,6364	98,1818	99,8	98,9471
22	Model V		99,6364	98,1818	99,8	98,9471
23	Model W		99,2744	96.2727	99,602	97.8451
24	Model X		99,0909	96.2626	98,6	97.3678



Gambar 4. 50 Grafik Perbandingan Kinerja Seluruh Model (A-X) berdasarkan Akurasi, Presisi, *Recall*, dan F1-Score.

Untuk mempermudah analisis komparatif dari Tabel 4.25, Gambar 4.50 mengilustrasikan perbandingan nilai rata-rata Akurasi, Presisi, *Recall*, dan F1-Score untuk kedua puluh empat konfigurasi model *Multilayer Perceptron* (Model A hingga X). Berdasarkan grafik tersebut, terlihat jelas bahwa secara keseluruhan model mampu memberikan performa klasifikasi yang sangat baik, di mana metrik Akurasi dan *Recall* menunjukkan tingkat kestabilan yang tinggi dengan nilai konsisten di atas 98%, yang membuktikan keandalan jaringan dalam meminimalkan kasus *false negative*. Di sisi lain, grafik ini juga memperlihatkan adanya fluktuasi yang lebih dinamis pada metrik Presisi dan F1-Score, dengan penurunan yang cukup kentara pada beberapa skenario seperti Model D, L, P, dan T. Fluktuasi ini menegaskan bahwa meskipun arsitektur jaringan sangat sensitif dalam mendeteksi kelas positif, pemilihan parameter spesifik seperti jumlah fitur dan variasi *hidden layer* memegang peranan krusial dalam menekan angka *false positive*, yang selanjutnya menjadi landasan dalam mengevaluasi dan menentukan konfigurasi model terbaik.

4.2.1 Evaluasi Variasi Parameter

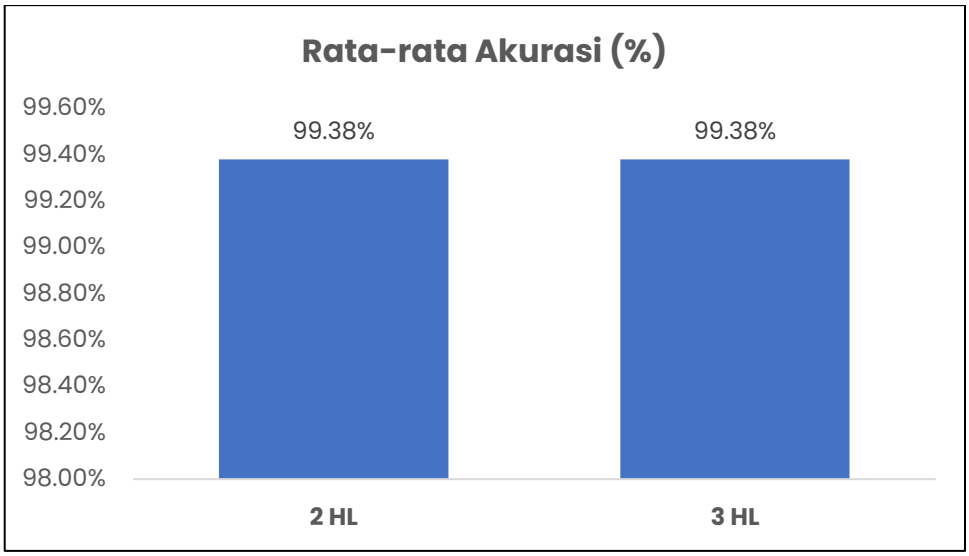
Subbab ini menjelaskan hasil evaluasi masing-masing parameter yang divariasikan dalam penelitian, yaitu jumlah *hidden layer*, konfigurasi neuron pada setiap lapisan tersembunyi, serta jumlah fitur input terhadap kinerja model MLP.

4.2.1.1 *Hidden Layer*

Perbandingan rata-rata akurasi antara kelompok arsitektur dengan 3 *hidden layer* dan 2 *hidden layer* disajikan pada Tabel 4.26.

Tabel 4. 26 Perbandingan Akurasi Rata-rata: 3 *Hidden Layer* vs 2 *Hidden Layer*

No	3 *Hidden Layer*		2 *Hidden Layer*	
	Model	**Akurasi(%)**	**Model**	**Akurasi(%)**
1	Model A	99,46	Model M	99,64
2	Model B	99,64	Model N	99,64
3	Model C	99,27	Model O	99,46
4	Model D	98,91	Model P	98,91
5	Model E	99,46	Model Q	99,64
6	Model F	99,64	Model R	99,64
7	Model G	99,46	Model S	99,28
8	Model H	99,28	Model T	98,73
9	Model I	99,64	Model U	99,64
10	Model J	99,46	Model V	99,64
11	Model K	99,46	Model W	99,27
12	Model L	98,91	Model X	99,10
	Rata-rata	**98,38**	**Rata-rata**	**99,38**



Gambar 4. 51 Perbandingan Rata-rata Akurasi

Hasil pengujian menunjukkan bahwa model dengan 3 *hidden layer* memiliki rata-rata akurasi 99,38%, sedangkan model dengan 2 *hidden layer* mencapai rata-rata 99,38%. Perbedaan rata-rata akurasi antara kedua kelompok sangat kecil, mengindikasikan bahwa pada dataset ekspresi gen LUSC yang telah melalui proses seleksi fitur yang memadai, kedalaman jaringan tidak memberikan dampak signifikan terhadap kinerja klasifikasi. Kondisi ini diduga disebabkan oleh karakteristik data ekspresi gen yang, setelah diseleksi menggunakan Mutual

Information, telah menghasilkan representasi fitur yang cukup informatif sehingga jaringan yang lebih dangkal pun mampu mempelajari pola yang relevan dengan baik. Temuan ini sejalan dengan prinsip parsimoni dalam pemodelan, di mana model yang lebih sederhana dengan performa setara lebih diutamakan karena memiliki risiko *overfitting* yang lebih rendah dan efisiensi komputasi yang lebih baik.

4.2.1.2 Neuron

Analisis kinerja berdasarkan konfigurasi neuron dilakukan secara terpisah untuk kelompok 3 *hidden layer* dan 2 *hidden layer*.

1) 3 *Hidden Layer*

Perbandingan akurasi model dengan arsitektur 3 *hidden layer* yang dibedakan berdasarkan jumlah neuron Tabel 4.27.

Tabel 4. 27 Perbandingan Akurasi Berdasarkan Konfigurasi Neuron pada 3 *Hidden Layer*

No	Neuron 64-32-16		Neuron 128–64–32		Neuron 256–128–64	
	Model	**Akurasi (*Mean*) (%)**	**Model**	**Akurasi (*Mean*) (%)**	**Model**	**Akurasi (*Mean*) (%)**
1	Model A	99,46	Model E	99,46	Model I	99,64
2	Model B	99,64	Model F	99,64	Model J	99,46
3	Model C	99,27	Model G	99,46	Model K	99,46
4	Model D	98,91	Model H	99,28	Model L	98,91
	Rata-rata	**99,32**	**Rata-rata**	**99,46**	**Rata-rata**	**99,37**

Dari Tabel 4.27 terlihat bahwa variasi konfigurasi neuron pada arsitektur 3 *hidden layer* menghasilkan perbedaan akurasi yang relatif kecil. Konfigurasi (256-128-64) menghasilkan rata-rata akurasi tertinggi sebesar 99,37%, diikuti konfigurasi (128-64-32) sebesar 99,46%, dan konfigurasi (64-32-16) sebesar 99,32%. Temuan ini menunjukkan bahwa arsitektur dengan kapasitas menengah hingga besar tidak memberikan peningkatan yang berarti, mengonfirmasi bahwa

fitur-fitur terpilih hasil *Mutual Information* sudah cukup informatif sehingga jaringan berukuran kompak pun mampu mengekstraksi pola yang relevan.

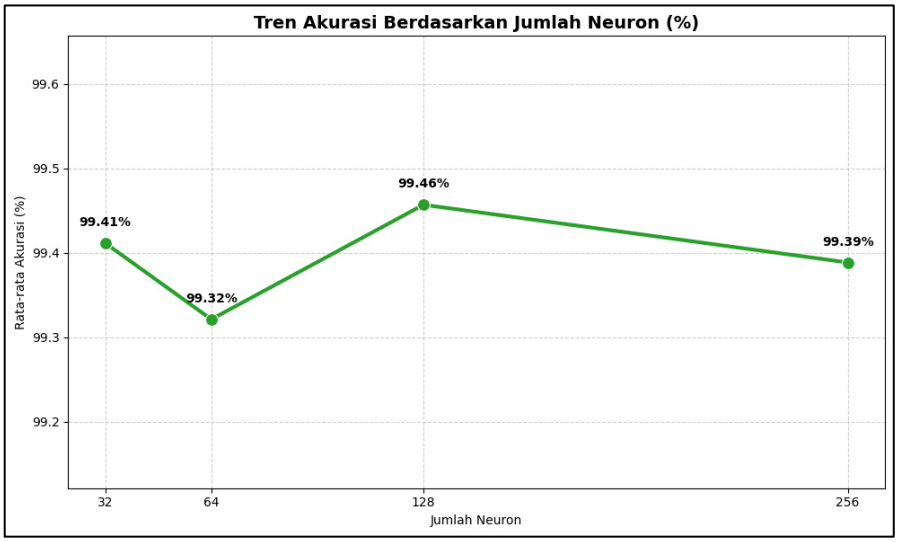
2) 2 *Hidden Layer*

Perbandingan akurasi model dengan arsitektur 2 *hidden layer* berdasarkan jumlah neuron disajikan pada Tabel 4.28.

Tabel 4. 28 Perbandingan Akurasi Berdasarkan Konfigurasi Neuron pada 2 *Hidden Layer*

No	Neuron 32-16		Neuron 64–32		Neuron 256–128	
	Model	**Akurasi (Mean) (%)**	**Model**	**Akurasi (Mean) (%)**	**Model**	**Akurasi (Mean) (%)**
1	Model M	99,64	Model Q	99,64	Model U	99,64
2	Model N	99,64	Model R	99,64	Model V	99,64
3	Model O	99,46	Model S	99,28	Model W	99,27
4	Model P	98,91	Model T	98,73	Model X	99,10
	Rata-rata	**99,41**	**Rata-rata**	**99,32**	**Rata-rata**	**99,41**

Dari Tabel 4.28 terlihat bahwa pada arsitektur 2 *hidden layer*, konfigurasi (256-128) menghasilkan rata-rata akurasi tertinggi sebesar 99,41%, diikuti (64-32) sebesar 99,32%, dan (32-16) sebesar 99,41%. Menariknya, arsitektur yang paling kompak, yaitu konfigurasi (32-16), tampil sangat kompetitif dan bahkan setara dengan konfigurasi yang lebih besar pada nilai fitur input yang rendah. Hal ini konsisten dengan prinsip Occam's Razor dalam *machine learning*, bahwa model yang lebih sederhana lebih diutamakan apabila mampu memberikan performa yang setara.



Gambar 4. 52 Pengaruh Jumlah Neuron pada *Hidden Layer* terhadap Rata-rata Akurasi.

Untuk menganalisis dampak variasi kelebaran jaringan (*width*), Gambar 4.52 menyajikan visualisasi perbandingan kinerja model berdasarkan variasi jumlah neuron. Berdasarkan hasil pengujian, peningkatan jumlah neuron hingga titik tertentu terbukti mampu meningkatkan kapasitas jaringan dalam mengekstraksi pola *gene expression*, di mana performa paling optimal dicapai pada konfigurasi 128 neuron dengan rata-rata akurasi sebesar 99,46%. Performa yang sangat kompetitif juga ditunjukkan oleh arsitektur yang lebih ringkas dengan 32 neuron yang mencapai akurasi 99,41%, sementara penggunaan 64 neuron menghasilkan akurasi sebesar 99,32%. Namun, ketika jumlah neuron ditingkatkan hingga 256 neuron, akurasi model justru mengalami penurunan menjadi 99,39%. Fenomena ini mengindikasikan bahwa kapasitas parameter yang terlalu masif (berlebihan) tidak berbanding lurus dengan peningkatan performa, melainkan rentan memicu *overfitting* di mana model mulai menghafal *noise* pada data latih alih-alih mempelajari pola karsinogenesis yang esensial. Oleh karena itu, penggunaan 128 neuron ditetapkan sebagai konfigurasi *width* yang paling proporsional untuk

menyeimbangkan representasi fitur yang kuat dengan kemampuan generalisasi model *Multilayer Perceptron*.

4.2.1.3 Input *Layer*

Subbab ini membahas evaluasi kinerja model deteksi berdasarkan jumlah fitur gen terpilih.

1) 3 *Hidden Layer*

Perbandingan akurasi model dengan arsitektur 3 *hidden layer* berdasarkan jumlah fitur input disajikan pada Tabel 4.25.

Tabel 4. 29 Perbandingan Akurasi Berdasarkan Jumlah Fitur pada 3 *Hidden Layer*

No	**3 *Hidden Layer***							
	50 Input		**100 Input**		**500 Input**		**1500 Input**	
	Model	**Akurasi (Mean)**	**Model**	**Akurasi (Mean)**	**Model**	**Akurasi (Mean)**	**Model**	**Akurasi (Mean)**
1	Model A	99,46	Model B	99,64	Model C	99,27	Model D	98,91
2	Model E	99,46	Model F	99,64	Model G	99,46	Model H	99,28
3	Model I	99,64	Model J	99,46	Model K	99,46	Model L	98,91
	Rata-rata	**99,52**	**Rata-rata**	**99,58**	**Rata-rata**	**99,40**	**Rata-rata**	**99,03**

Dari Tabel 4.29 terlihat bahwa pada arsitektur 3 *hidden layer*, penggunaan 100 fitur menghasilkan rata-rata akurasi tertinggi sebesar 99,58%, diikuti 50 fitur sebesar 99,52%, 500 fitur sebesar 99,40%, dan 1500 fitur sebesar 99,03%. Pola ini mengindikasikan adanya titik optimal penggunaan fitur pada n_MI = 100, Penggunaan fitur yang terlalu sedikit (50 gen) menyebabkan representasi informasi yang kurang lengkap, sedangkan penggunaan fitur yang terlalu banyak (1500 gen) memperkenalkan noise dari gen-gen yang kurang relevan, sehingga menurunkan performa sekaligus meningkatkan ketidakstabilan model yang tercermin dari nilai standar deviasi yang lebih besar.

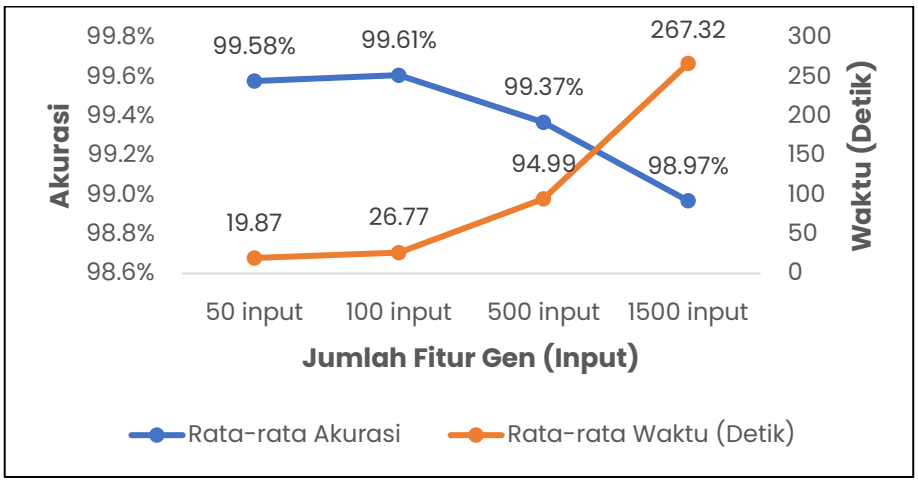
2) 2 *Hidden Layer*

Perbandingan akurasi model dengan arsitektur 2 *hidden layer* berdasarkan jumlah fitur input disajikan pada Tabel 4.30,

Tabel 4. 30 Perbandingan Akurasi Berdasarkan Jumlah Fitur pada 2 *Hidden Layer*

No	***2 Hidden Layer***							
	50 Input		**100 Input**		**500 Input**		**1500 Input**	
	Model	**Akurasi (*Mean*)**	**Model**	**Akurasi**	**Model**	**Akurasi**	**Model**	**Akurasi**
1	Model M	99,64	Model N	99,64	Model O	99,46	Model P	98,91
2	Model Q	99,64	Model R	99,64	Model S	99,28	Model T	98,73
3	Model U	99,64	Model V	99,64	Model W	99,27	Model X	99,10
	Rata-rata	**99,64**	**Rata-rata**	**99,64**	**Rata-rata**	**99,34**	**Rata-rata**	**98,91**

Dari Tabel 4.30 terlihat bahwa pada arsitektur 2 *hidden layer*, penggunaan 50 fitur menghasilkan rata-rata akurasi tertinggi sebesar 99,64%, diikuti 100 fitur sebesar 99,64%, 500 fitur sebesar 99,34%, dan 1500 fitur sebesar 98,91%. Pola ini sedikit berbeda dari arsitektur 3 *hidden layer*, yang menunjukkan bahwa interaksi antara kapasitas jaringan dan jumlah fitur bersifat saling memengaruhi. Secara keseluruhan, tren penurunan akurasi seiring bertambahnya jumlah fitur melampaui titik optimal mengonfirmasi adanya efek curse of dimensionality, yaitu kondisi di mana penambahan fitur yang berlebihan justru mempersulit proses pembelajaran akibat meningkatnya rasio noise terhadap sinyal yang bermakna.



Gambar 4. 53 Pengaruh Jumlah Fitur Gen (Input) terhadap Rata-rata Akurasi dan Waktu Komputasi.

Untuk memperjelas dampak dari reduksi dimensi yang dilakukan, Gambar 4.53 memvisualisasikan *trade-off* antara rata-rata akurasi keseluruhan model (garis biru) dengan rata-rata waktu komputasi (garis oranye). Berdasarkan grafik tersebut, penggunaan 100 fitur gen terbukti menjadi titik optimal (*sweet spot*) yang menghasilkan akurasi tertinggi sebesar 99,61% dengan waktu pelatihan yang sangat efisien, yakni hanya 26,77 detik. Penggunaan fitur yang lebih sedikit (50 input) menyebabkan sedikit penurunan akurasi menjadi 99,58% akibat berkurangnya informasi biologis yang relevan. Sebaliknya, penambahan fitur secara berlebihan hingga 1500 input tidak hanya menurunkan akurasi secara signifikan menjadi 98,97% akibat fenomena *curse of dimensionality* (di mana gen-gen yang kurang relevan bertindak sebagai *noise*), tetapi juga memicu lonjakan beban komputasi secara drastis hingga mencapai 267,32 detik. Hal ini menegaskan bahwa reduksi dimensi menggunakan metode *Mutual Information* pada 100 fitur teratas adalah langkah yang sangat krusial dan efektif untuk memaksimalkan akurasi deteksi model MLP sekaligus menjaga efisiensi komputasi.

4.2.2 Analisis K-*Fold Cross validation* pada Model Terbaik

K-*Fold Cross validation* dengan k=5 diterapkan pada seluruh model untuk memastikan evaluasi performa tidak bergantung pada satu skenario pembagian data tertentu. Metode ini membagi dataset menjadi lima bagian, di mana model dilatih pada empat bagian dan diuji pada satu bagian secara bergantian. Pendekatan ini memungkinkan penilaian kestabilan dan keandalan model secara lebih objektif. Subbab berikut menganalisis kestabilan performa model terbaik pada kelompok 3 *hidden layer* dan 2 *hidden layer*.

4.2.2.1 Model Terbaik 3 *Hidden layer*

Berdasarkan hasil evaluasi terhadap seluruh dua belas konfigurasi arsitektur tiga *hidden layer* (Model A hingga Model L) yang telah dirangkum pada Tabel 4.21 dan Tabel 4.22, Model B ditetapkan sebagai konfigurasi terbaik dalam kelompok ini. Model B menggunakan arsitektur tiga *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan memanfaatkan 100 fitur terpilih hasil seleksi berbasis *Mutual Information* (MI). Penetapan Model B sebagai model terbaik didasarkan pada pertimbangan nilai metrik *Recall* dan F1-Score tertinggi yang dicapainya. Model B berhasil memperoleh nilai rata-rata *Recall* sebesar 99,80% (σ=0,0024) dan F1-Score sebesar 98,95% (σ=0,0129). Nilai σ *Recall* sebesar 0,0024 merupakan salah satu yang terkecil di antara seluruh konfigurasi tiga *hidden layer*, mengindikasikan bahwa kemampuan deteksi Model B tidak hanya unggul secara rata-rata, tetapi juga bersifat sangat konsisten dan stabil di antara kelima *fold* pengujian. Keunggulan pada metrik *Recall* secara khusus menjadi pertimbangan utama mengingat konteks

aplikasi klasifikasi kanker paru-paru, di mana meminimalkan jumlah *false negative* atau kasus kanker yang tidak terdeteksi merupakan prioritas klinis yang kritis.

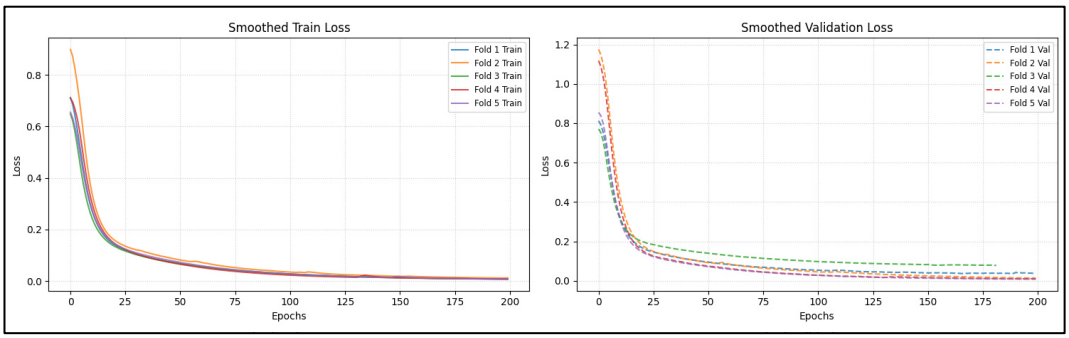
4.2.2.2 Model Terbaik 2 Hidden layer.

Berdasarkan hasil evaluasi terhadap seluruh dua belas konfigurasi arsitektur dua *hidden layer* (Model M hingga Model X) yang telah dirangkum pada Tabel 4.21 dan Tabel 4.22, Model M ditetapkan sebagai konfigurasi terbaik dalam kelompok ini. Model M dikonfigurasi menggunakan arsitektur dua *hidden layer* yang lebih dangkal dengan konfigurasi neuron (32, 16) dan hanya memanfaatkan 50 fitur input yang dipilih melalui seleksi berbasis *Mutual Information* (MI). Meskipun menggunakan jumlah fitur yang lebih sedikit dan arsitektur yang lebih ringkas dibandingkan sejumlah konfigurasi dua *hidden layer* lainnya, Model M mampu mencapai nilai rata-rata *Recall* sebesar 99,80% (σ=0,0024) dan F1-Score sebesar 98,95% (σ=0,0129). Nilai σ *Recall* yang sangat kecil yakni 0,0024 mencerminkan kestabilan kemampuan deteksi yang luar biasa tinggi dan konsisten di antara seluruh partisi data pengujian. Capaian ini mengindikasikan bahwa arsitektur yang lebih sederhana tidak menurunkan performa secara signifikan, dan bahwa 50 fitur terpilih telah cukup merepresentasikan informasi biologis yang relevan untuk membedakan sampel Tumor dari sampel Normal pada dataset ekspresi gen kanker paru LUSC.

4.2.2.3 Analisis Loss Curve Model Terbaik

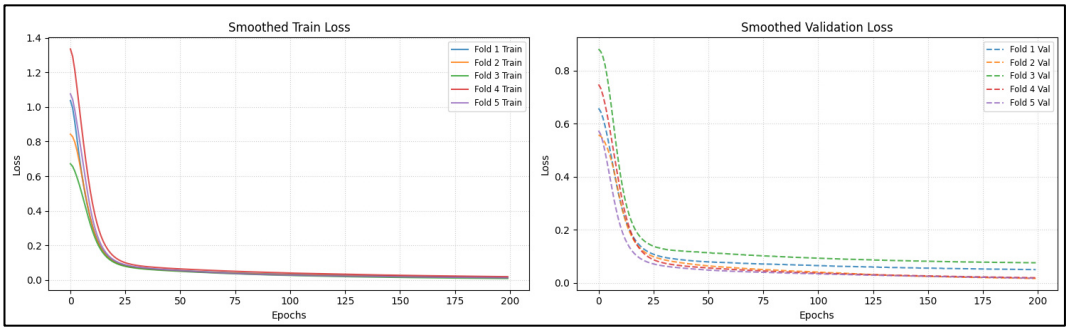
Untuk memverifikasi bahwa model terbaik tidak mengalami *overfitting* selama proses pelatihan, dilakukan analisis kurva *loss* pada Model B dan Model M

menggunakan data dari semua *fold* . Kurva training *loss* dan validation *loss* untuk kedua model disajikan pada Gambar 4.50 dan 4.51.



Gambar 4. 54 Kurva Training *Loss* dan Validation *Loss* Model B

Berdasarkan Gambar 4.54, Model B menunjukkan pola konvergensi yang sangat baik. Pada epoch-epoch awal, training *loss* dan validation *loss* berfluktuasi relatif tinggi, kemudian keduanya turun secara konsisten dan beriringan menuju nilai yang mendekati nol. Tidak ditemukan pola divergensi, yaitu kondisi di mana validation *loss* mulai meningkat sementara training *loss* terus menurun, yang merupakan tanda utama *overfitting*. Mekanisme *Early Stopping* terbukti efektif dalam menghentikan pelatihan pada kondisi generalisasi yang optimal.



Gambar 4. 55 Kurva Training *Loss* dan Validation *Loss* Model M

Pola serupa juga terlihat pada Model M dalam Gambar 4.55. Training *loss* awal turun secara konsisten bersamaan dengan validation *loss*. Meskipun pada

epoch awal validation *loss* sempat lebih rendah dibandingkan training *loss*, hal ini disebabkan oleh penerapan SMOTE hanya pada data latih sehingga distribusinya lebih bervariasi. Setelah beberapa epoch, keduanya secara bertahap mendekati nilai nol dan mempertahankannya hingga mekanisme *Early Stopping* aktif. Gap antara training *loss* dan validation *loss* tidak melebar sepanjang proses pelatihan, mengonfirmasi bahwa Model M mampu menggeneralisasi dengan baik. Dengan demikian, analisis kurva *loss* pada kedua model terbaik mengindikasikan bahwa proses pelatihan berjalan secara optimal tanpa tanda-tanda *overfitting*.

4.2.2.4 Pemilihan Model Terbaik Keseluruhan

Tahap akhir analisis dilakukan dengan membandingkan kedua model terbaik dari masing-masing kelompok arsitektur, yakni Model B yang mewakili konfigurasi tiga *hidden layer* terbaik dan Model M yang mewakili konfigurasi dua *hidden layer* terbaik. Merujuk pada visualisasi evaluasi sebelumnya yang menyajikan distribusi performa kedua model secara keseluruhan melalui *boxplot*, terlihat bahwa kedua model memperlihatkan karakteristik distribusi yang sangat serupa pada seluruh metrik evaluasi.

Secara statistik, Model B dan Model M menghasilkan performa prediksi yang identik. Kedua model memperoleh nilai rata-rata *Recall* sebesar 99,80% dengan standar deviasi σ=0,0024 dan nilai rata-rata F1-Score sebesar 98,95% dengan standar deviasi σ=0,0129. Kesamaan nilai rata-rata dan standar deviasi yang persis sama pada kedua metrik kunci tersebut mengindikasikan bahwa secara statistik tidak terdapat perbedaan performa yang bermakna di antara keduanya.

Menghadapi kondisi performa yang setara secara statistik antara dua kandidat model, pemilihan model terbaik akhir didasarkan pada Prinsip *Occam's Razor* atau hukum parsimoni. Prinsip ini menyatakan bahwa apabila dua model berkinerja setara, maka model yang lebih sederhana yakni model dengan jumlah parameter dan dimensi ruang model yang lebih kecil seharusnya lebih diutamakan dibandingkan model yang lebih kompleks (Bargagli Stoffi et al., 2022). Lebih lanjut, keunggulan model yang lebih kompleks umumnya diimbangi oleh sejumlah kerugian nyata, di antaranya kompleksitas implementasi yang lebih tinggi dan kebutuhan sumber daya komputasi yang lebih besar, sehingga dalam praktiknya model yang lebih sederhana tetap lebih disukai selama performanya setara.

Berdasarkan argumen tersebut, Model M ditetapkan sebagai model terbaik keseluruhan (overall best model) dalam penelitian ini. Dibandingkan Model B yang menggunakan arsitektur tiga *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan 100 fitur input, Model M mengandalkan arsitektur dua *hidden layer* yang lebih dangkal dengan konfigurasi neuron (32, 16) dan hanya 50 fitur input, sehingga merupakan solusi yang lebih parsimonious. Penetapan ini juga didukung oleh pertimbangan kepraktisan deployment, di mana arsitektur yang lebih sederhana membutuhkan sumber daya komputasi yang lebih rendah dan waktu inferensi yang lebih singkat pada sistem klinis nyata, tanpa mengorbankan sedikit pun kemampuan deteksi yang diharapkan.

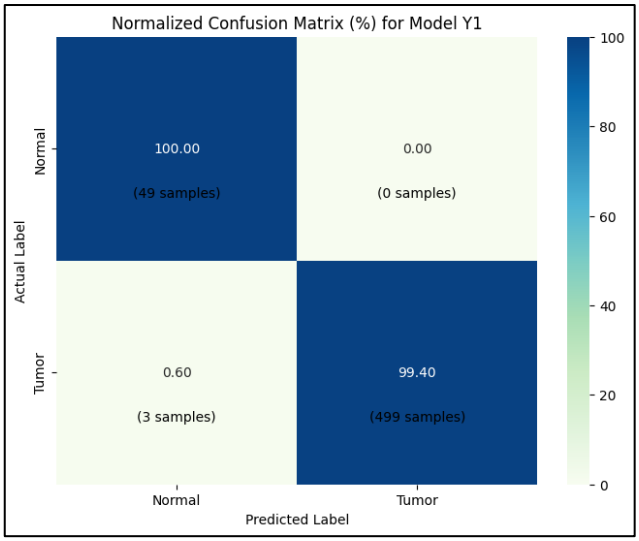
4.3 Perbandingan Variasi Nilai K pada Model Terbaik

Untuk menganalisis pengaruh variasi jumlah *fold* terhadap kestabilan performa model terbaik, dilakukan pengujian tambahan menggunakan skema 3-

Fold dan 10-*Fold Cross validation* pada konfigurasi arsitektur dua *hidden layer* terbaik (Model M) dan tiga *hidden layer* terbaik (Model B). Pengujian ini menghasilkan empat skenario tambahan, yaitu Model Y1, Y2, Z1, dan Z2.

4.2.3.1 Model Y1

Model Y1 dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI), identik dengan konfigurasi Model M, namun diuji menggunakan skema 3-*Fold Cross validation* (K=3) untuk menganalisis sensitivitas performa terhadap variasi jumlah *fold* pengujian. Model ini tetap memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 dan mekanisme *Early Stopping* yang sama dengan konfigurasi sebelumnya.

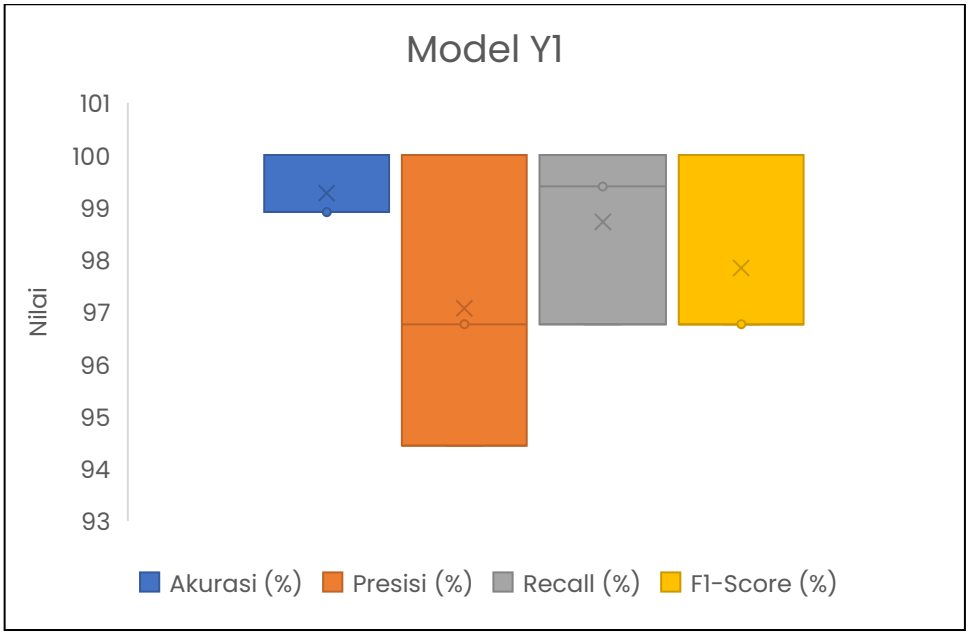


Gambar 4. 56 Metrik Evaluasi Model Y1

Gambar 4. 57 Metrik Evaluasi Model Y1

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	99,46	97.22	99,70	98,42
Fold 2	98,91	94.44	99,40	96.76
Fold 3	100,00	100,00	100,00	100,00
Rata-rata	**99,46**	**97.22**	**99,70**	**98,39**

Berdasarkan confusion *matrix* pada Gambar 4.52, Model Y1 mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* tanpa satu pun *false positive*, serta 499 dari 502 data kanker sebagai *true positive*, dengan 3 data (0,60%) yang terklasifikasi keliru sebagai *false negative*. Secara keseluruhan, model mencapai rata-rata akurasi 99,46%, presisi 97.22%, *recall* 99,70%, dan F1-score 98,39%. Nilai standar deviasi akurasi sebesar 0,0045 menunjukkan konsistensi performa yang sangat baik di antara ketiga *fold* pengujian.



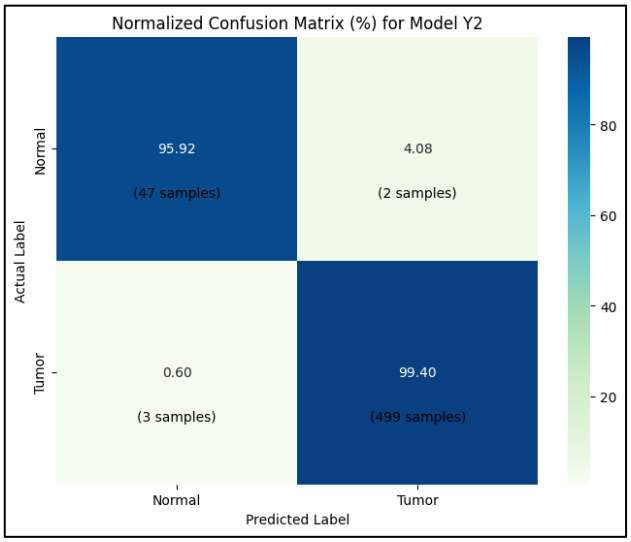
Gambar 4. 58 Visualisasi Rata-rata evaluation *matrix* Model Y1

Distribusi performa Model Y1 divisualisasikan melalui grafik *boxplot*. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0045), Presisi (*Macro*) μ=0,9722 (σ=0,0227), *Recall* (*Macro*) μ=0,9970 (σ=0,0024), dan F1 (*Macro*) μ=0,9839 (σ=0,0132). Pola distribusi ini sangat serupa dengan Model M pada skema 5-*Fold* , di mana kotak IQR Akurasi dan *Recall* (*Macro*) tampak sempit

di rentang nilai tinggi yang mencerminkan kestabilan deteksi yang sangat baik, sementara kotak IQR Presisi (*Macro*) tetap menunjukkan sebaran yang lebih lebar. Hal ini mengindikasikan bahwa penurunan jumlah *fold* dari 5 menjadi 3 tidak memberikan dampak signifikan terhadap kestabilan performa model.

4.2.3.2 Model Y2

Model Y2 dikonfigurasi menggunakan 2 *hidden layer* dengan konfigurasi neuron (32, 16) dan jumlah fitur input sebesar 50 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI), identik dengan konfigurasi Model M, namun diuji menggunakan skema 10-*Fold Cross validation* (K=10) untuk menganalisis sensitivitas performa terhadap variasi jumlah *fold* pengujian. Model ini tetap memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 dan mekanisme *Early Stopping* yang sama dengan konfigurasi sebelumnya.



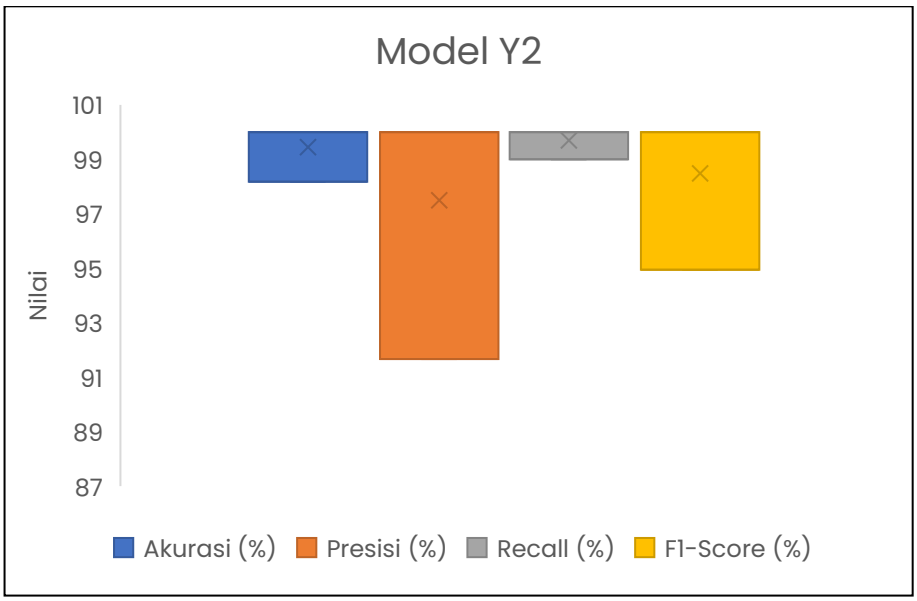
Gambar 4. 59 Metrik Evaluasi Model Y2

Tabel 4. 31 Metrik Evaluasi Model Y2

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,21	99,04	90,00	93.96
Fold 2	96.36	89.00	89.00	89.00

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 3	100,00	100,00	100,00	100,00
Fold 4	98,18	91.67	99,00	94.95
Fold 5	98,18	91.67	99,00	94.95
Fold 6	100,00	100,00	100,00	100,00
Fold 7	100,00	100,00	100,00	100,00
Fold 8	100,00	100,00	100,00	100,00
Fold 9	100,00	100,00	100,00	100,00
Fold 10	100,00	100,00	100,00	100,00
Rata-rata	**99,09**	**97.14**	**97.70**	**97.29**

Berdasarkan confusion *matrix* yang dihasilkan, Model Y2 mendeteksi 47 dari 49 data jaringan normal sebagai *true negative* (95.92%), dengan 2 data (4.08%) yang terklasifikasi keliru sebagai *false positive*, serta 499 dari 502 data kanker sebagai *true positive* (99,40%), dengan 3 data (0,60%) yang terklasifikasi keliru sebagai *false negative*. Secara keseluruhan, model mencapai rata-rata akurasi 99,09%, presisi 97.14%, *recall* 97.70%, dan F1-score 97.29%. Nilai standar deviasi akurasi sebesar 0,0122 menunjukkan fluktuasi performa yang lebih besar dibandingkan skema 3-*Fold* maupun 5-*Fold* , terutama dipengaruhi oleh hasil *Fold* 1 dan *Fold* 2 yang menunjukkan penurunan *recall* cukup tajam hingga 89,00–90,00%.



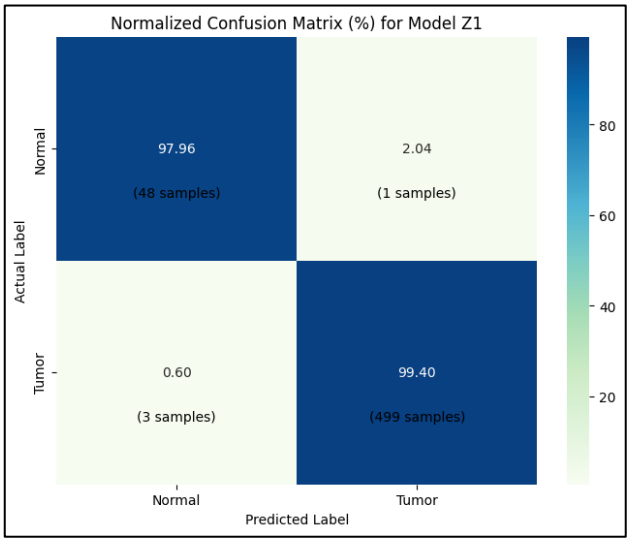
Gambar 4. 60 Visualisasi Rata-rata evaluation *matrix* Model Y2

Distribusi performa Model Y2 divisualisasikan melalui grafik *boxplot*. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9909 (σ=0,0122), Presisi (*Macro*) μ=0,9714 (σ=0,0423), *Recall* (*Macro*) μ=0,9770 (σ=0,0412), dan F1 (*Macro*) μ=0,9728 (σ=0,0367). Dibandingkan dengan Model Y1, seluruh metrik pada Model Y2 menunjukkan nilai standar deviasi yang jauh lebih besar, terutama pada *Recall* (*Macro*) yang meningkat tajam dari σ=0,0024 menjadi σ=0,0412. Hal ini mengindikasikan bahwa peningkatan jumlah *fold* menjadi 10 menyebabkan ukuran data uji pada setiap *fold* menjadi lebih kecil, sehingga fluktuasi performa antar *fold* menjadi lebih besar dan kestabilan model relatif menurun dibandingkan skema 3-*Fold* maupun 5-*Fold* .

4.2.3.3 Model Z1

Model Z1 dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui

seleksi berbasis *Mutual Information* (MI), identik dengan konfigurasi Model B, namun diuji menggunakan skema 3-*Fold Cross validation* (K=3) untuk menganalisis sensitivitas performa terhadap variasi jumlah *fold* pengujian. Model ini tetap memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 dan mekanisme *Early Stopping* yang sama dengan konfigurasi sebelumnya.



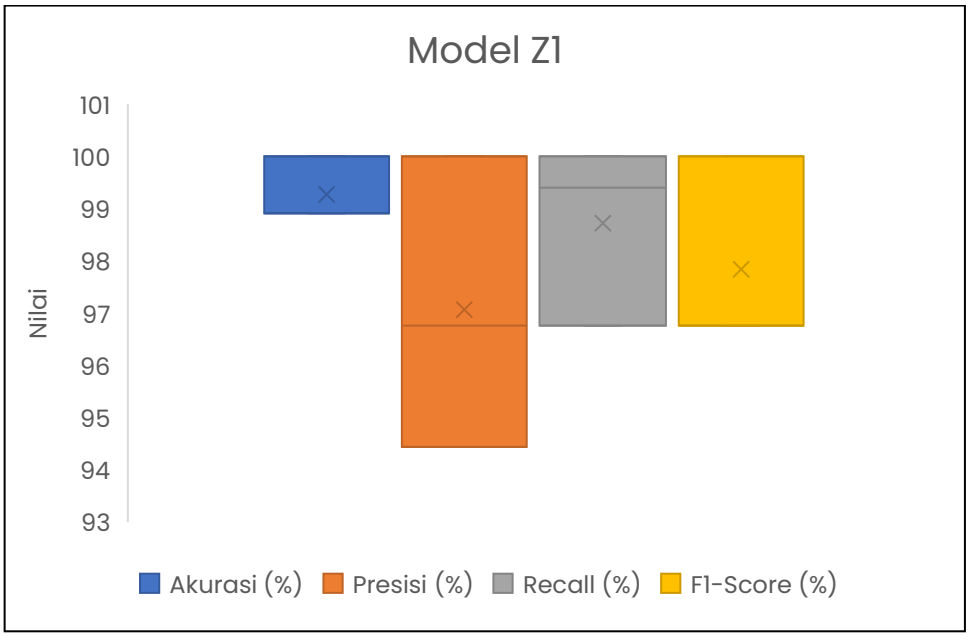
Gambar 4. 61 Metrik Evaluasi Model Z1

Tabel 4. 32 Metrik Evaluasi Model Z1

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	98,91	96.76	96.76	96.76
Fold 2	98,91	94.44	99,40	96.76
Fold 3	100,00	100,00	100,00	100,00
Rata-rata	**99,28**	**97.07**	**98,72**	**97.84**

Berdasarkan confusion *matrix* yang dihasilkan, Model Z1 mendeteksi 48 dari 49 data jaringan normal sebagai *true negative* (97,96%), dengan 1 data (2,04%) yang terklasifikasi keliru sebagai *false positive*, serta 499 dari 502 data kanker sebagai *true positive* (99,40%), dengan 3 data (0,60%) yang terklasifikasi keliru sebagai *false negative*. Secara keseluruhan, model mencapai rata-rata akurasi 99,28%, presisi 97,07%, *recall* 98,72%, dan F1-score 97,84%. Nilai standar deviasi

akurasi sebesar 0,0052 menunjukkan konsistensi performa yang baik di antara ketiga *fold* pengujian.



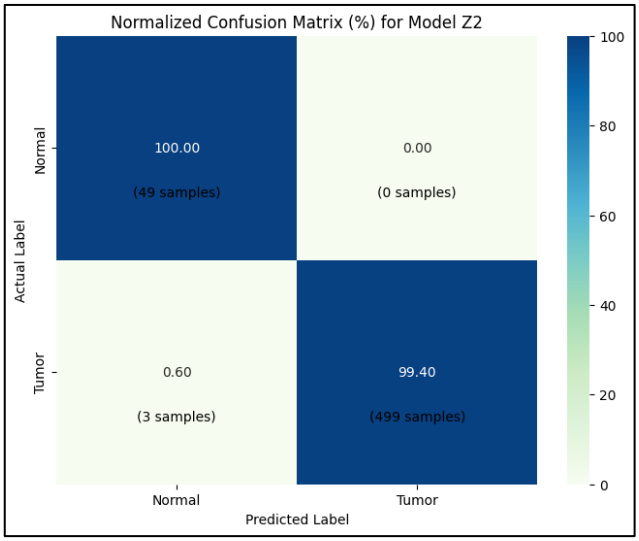
Gambar 4. 62 Visualisasi Rata-rata evaluation *matrix* Model Z1

Distribusi performa Model Z1 divisualisasikan melalui grafik *boxplot*. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9927 (σ=0,0052), Presisi (*Macro*) μ=0,9707 (σ=0,0228), *Recall* (*Macro*) μ=0,9872 (σ=0,0141), dan F1 (*Macro*) μ=0,9784 (σ=0,0153). Dibandingkan dengan Model B pada skema 5-*Fold* , nilai rata-rata *Recall* pada Model Z1 sedikit menurun (98,72% berbanding 99,80%) disertai standar deviasi yang sedikit lebih besar, namun secara keseluruhan model masih menunjukkan kestabilan yang baik pada metrik Akurasi dan *Recall*.

4.2.3.4 Model Z2

Model Z2 dikonfigurasi menggunakan 3 *hidden layer* dengan konfigurasi neuron (64, 32, 16) dan jumlah fitur input sebesar 100 fitur yang dipilih melalui seleksi berbasis *Mutual Information* (MI), identik dengan konfigurasi Model B,

namun diuji menggunakan skema 10-*Fold Cross validation* (K=10) untuk menganalisis sensitivitas performa terhadap variasi jumlah *fold* pengujian. Model ini tetap memanfaatkan teknik regularisasi L2 dengan *lambda* 0,001 dan mekanisme *Early Stopping* yang sama dengan konfigurasi sebelumnya.



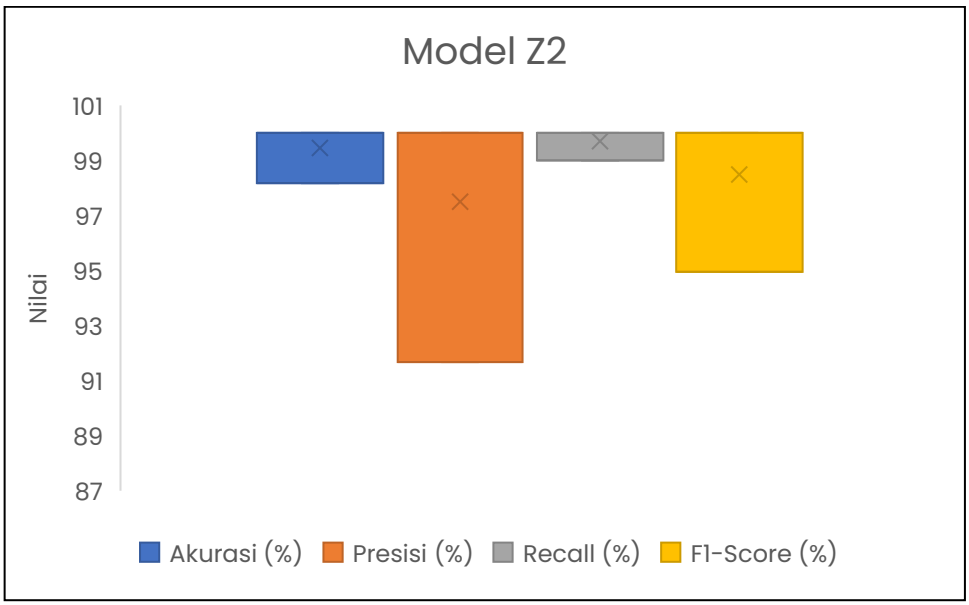
Gambar 4. 63 Metrik Evaluasi Model Z2

Tabel 4. 33 Metrik Evaluasi Model Z2

Fold	**Akurasi (%)**	**Presisi (%)**	***Recall* (%)**	**F1-Score (%)**
Fold 1	100,00	100,00	100,00	100,00
Fold 2	98,18	91.67	99,00	94.95
Fold 3	100,00	100,00	100,00	100,00
Fold 4	98,18	91.67	99,00	94.95
Fold 5	98,18	91.67	99,00	94.95
Fold 6	100,00	100,00	100,00	100,00
Fold 7	100,00	100,00	100,00	100,00
Fold 8	100,00	100,00	100,00	100,00
Fold 9	100,00	100,00	100,00	100,00
Fold 10	100,00	100,00	100,00	100,00
Rata-rata	**99,45**	**97.50**	**99,70**	**98,48**

Berdasarkan confusion *matrix* yang dihasilkan, Model Z2 mampu mendeteksi seluruh 49 data jaringan normal sebagai *true negative* tanpa satu pun *false positive*, serta 499 dari 502 data kanker sebagai *true positive* (99,40%), dengan 3 data (0,60%) yang terklasifikasi keliru sebagai *false negative*. Secara keseluruhan,

model mencapai rata-rata akurasi 99,45%, presisi 97.50%, *recall* 99,70%, dan F1-score 98,48%. Nilai standar deviasi akurasi sebesar 0,0083 menunjukkan performa yang relatif stabil meskipun terdapat tiga *fold* (*Fold* 2, 4, dan 5) yang menunjukkan penurunan presisi hingga 91,67%.

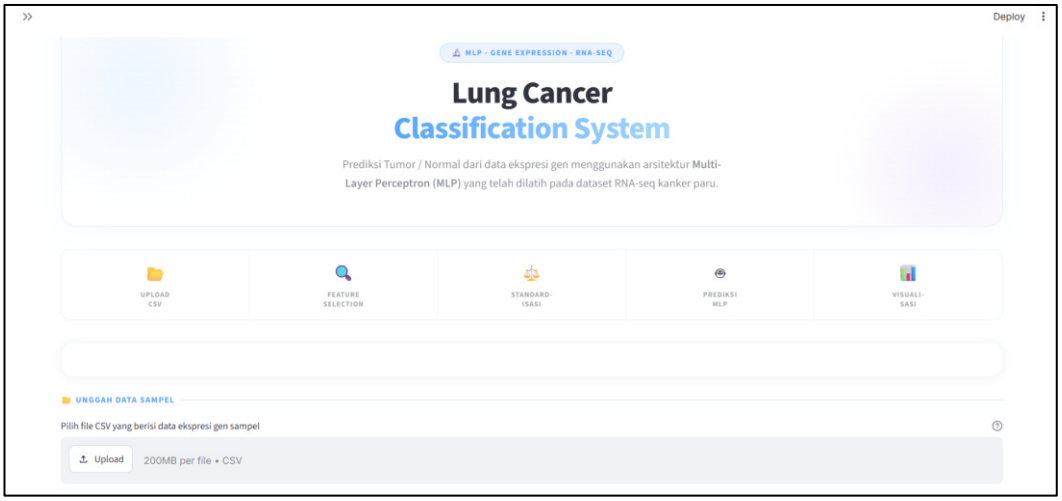


Gambar 4. 64 Visualisasi Rata-rata evaluation *matrix* Model Z2

Distribusi performa Model Z2 divisualisasikan melalui grafik *boxplot*. Berdasarkan grafik tersebut, diperoleh nilai Akurasi μ=0,9945 (σ=0,0083), Presisi (*Macro*) μ=0,9750 (σ=0,0382), *Recall* (*Macro*) μ=0,9970 (σ=0,0046), dan F1 (*Macro*) μ=0,9848 (σ=0,0231). Nilai rata-rata *Recall* dan F1-Score pada Model Z2 sangat mendekati capaian Model B pada skema 5-*Fold* , namun dengan standar deviasi Presisi dan F1 yang jauh lebih besar, mengindikasikan bahwa meskipun nilai tengah performa tetap tinggi, peningkatan jumlah *fold* menjadi 10 menyebabkan variabilitas antar *fold* yang lebih besar akibat berkurangnya ukuran data uji pada masing-masing *fold* .

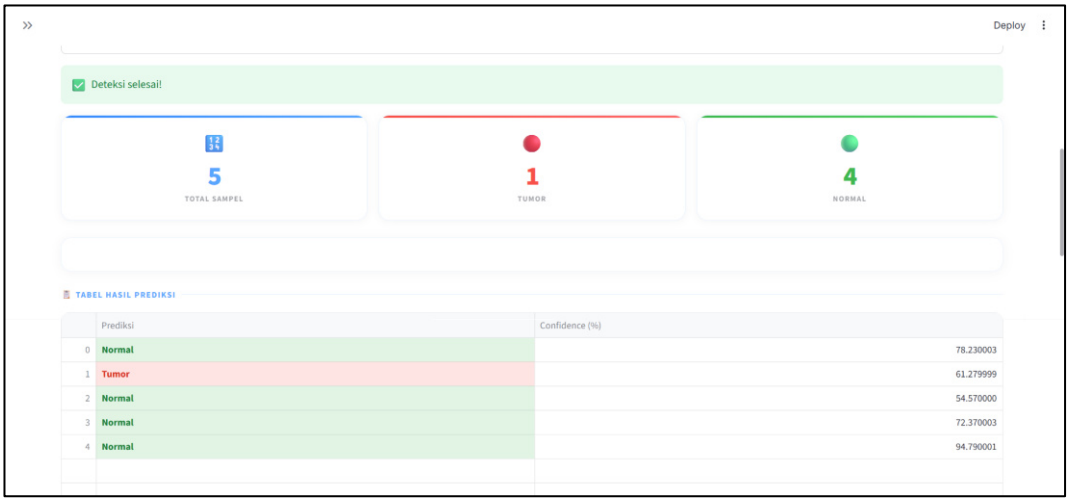
4.4 Implementasi GUI

Model klasifikasi kanker paru selanjutnya diaplikasikan dalam bentuk aplikasi berbasis web melalui perancangan *Graphical User Interface* (GUI) sebagai tahap akhir penelitian. GUI berfungsi sebagai media interaksi pengguna dalam menguji data ekspresi gen (*gene expression*) dan menampilkan hasil prediksi sistem. Aplikasi ini dikembangkan menggunakan *Streamlit*, yaitu *framework Python* yang memungkinkan integrasi model *machine learning* dengan antarmuka web secara sederhana dan efisien tanpa pengembangan front-end yang kompleks. Tampilan antarmuka utama sistem dapat dilihat pada Gambar 4.65.



Gambar 4. 65 GUI Sistem Klasifikasi Kanker Paru Berbasis Ekspresi Gen

Pengguna mengunggah file CSV yang berisi data sampel ekspresi gen (RNA-seq) melalui sistem, yang selanjutnya diproses melalui serangkaian tahapan pipeline meliputi seleksi fitur (*Feature Selection*) dan standardisasi data. Data yang telah diproses kemudian menjadi masukan bagi arsitektur *Multilayer Perceptron* (MLP) untuk menentukan label klasifikasi sampel.



Gambar 4. 66 Tampilan Hasil Klasifikasi Data Ekspresi Gen pada GUI

Gambar 4.66 menampilkan halaman hasil klasifikasi, di mana setiap data sampel ditampilkan di dalam tabel bersama label prediksi (Tumor atau Normal) serta nilai probabilitas (*Confidence*) dalam bentuk persentase. Perbedaan kelas prediksi ditampilkan secara visual dengan indikator warna yang berbeda untuk memudahkan interpretasi hasil. Selain itu, sistem juga menyediakan dashboard ringkasan di bagian atas yang menampilkan jumlah total sampel yang diuji serta distribusi jumlah sampel pada masing-masing kelas (Tumor dan Normal). Berdasarkan hasil implementasi, GUI yang dikembangkan mampu menjalankan fungsi sistem dengan baik dan konsisten dengan pipeline pengujian model, sehingga model klasifikasi deteksi kanker paru ini dapat diimplementasikan dalam bentuk aplikasi yang mudah digunakan.

Hasil penelitian ini sejalan dengan firman Allah SWT dalam QS. Al-Anbiya' [21]:30:

أَوَ لَمۡ يَرَ ٱلَّذِينَ كَفَرُوٓاْ أَنَّ ٱلسَّمَٰوَٰتِ وَٱلۡأَرۡضَ كَانَتَا رَتۡقٗا فَفَتَقۡنَٰهُمَاۖ وَجَعَلۡنَا مِنَ ٱلۡمَآءِ كُلَّ شَيۡءٍ حَيٍّۖ أَفَلَا يُؤۡمِنُونَ

٣٠

"Dan apakah orang-orang kafir tidak mengetahui bahwa langit dan bumi keduanya dahulunya menyatu, kemudian Kami pisahkan antara keduanya; dan Kami jadikan segala sesuatu yang hidup berasal dari air; maka mengapa mereka tidak beriman?" (QS. Al-Anbiya':30)

Menurut tafsir Ibnu Katsir, ayat ini mendorong manusia untuk mengamati dan merenungkan ciptaan Allah SWT, termasuk pada tingkat yang paling kecil sekalipun seperti sistem biologis dalam tubuh manusia (Al Shaykh & Abdul Ghoffar, 2006). Data ekspresi gen merupakan salah satu wujud tanda kekuasaan Allah pada ciptaan-Nya yang paling kompleks, yaitu ribuan gen yang bekerja secara teratur dalam setiap sel. Penelitian ini, dengan memanfaatkan kecerdasan buatan untuk membaca pola tersembunyi dalam data genomik, merupakan bentuk nyata dari perintah untuk menelaah ciptaan Allah demi kemaslahatan manusia. Kemampuan model MLP dalam mengklasifikasikan jaringan normal dan kanker dengan akurasi yang sangat tinggi mencerminkan betapa teraturnya sistem biologis yang Allah ciptakan, sekaligus membuka peluang bagi deteksi dini kanker paru berbasis genomik yang dapat meningkatkan kualitas hidup manusia.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil pengujian dan analisis mendalam menggunakan metode *Multilayer Perceptron* (MLP) yang dikombinasikan dengan seleksi fitur *Mutual Information* (MI) pada dataset TCGA-LUSC, penelitian ini berhasil menjawab rumusan masalah mengenai performa klasifikasi dalam membedakan jaringan paru normal dan jaringan kanker *Lung Squamous Cell Carcinoma* (LUSC). Hasil eksperimen membuktikan secara empiris bahwa performa klasifikasi yang optimal tidak selalu membutuhkan arsitektur jaringan saraf yang dalam dan kompleks, melainkan sangat bergantung pada reduksi dimensi yang tepat guna menghindari efek *curse of dimensionality*.

Hal ini ditunjukkan oleh keberhasilan Model M sebagai model terbaik keseluruhan yang memegang prinsip parsimoni; dengan hanya menggunakan arsitektur ringkas 2 *hidden layer* (32 dan 16 neuron) serta konfigurasi input minimal 50 fitur gen terbaik, model ini mampu mencapai rata-rata akurasi sebesar 99,64%, presisi 99,67%, *recall* 99,64%, dan F1-score 99,64%. Tingkat konsistensi yang sangat tinggi di antara lima *fold* pengujian (standar deviasi akurasi sebesar 0,0044) membuktikan bahwa model ini sangat stabil dan andal. Dengan demikian, penerapan model MLP dengan arsitektur kompak dan seleksi fitur berbasis MI terbukti menjadi pendekatan komputasi yang sangat efektif, akurat, sekaligus efisien untuk deteksi dini kanker paru berbasis data genomik.

5.2 Saran

Berdasarkan hasil dan keterbatasan penelitian ini, beberapa saran dapat dikemukakan untuk pengembangan lebih lanjut:

1. Penelitian ini menggunakan dataset TCGA-LUSC dengan jumlah sampel yang terbatas, khususnya pada kelas normal yang hanya terdiri dari 49 sampel. Ketidakseimbangan kelas yang signifikan ini, meskipun telah diatasi dengan teknik SMOTE, berpotensi membatasi kemampuan generalisasi model. Oleh karena itu, penambahan data dari sumber lain seperti GEO (*Gene expression* Omnibus) atau integrasi dengan dataset penelitian *Multi*-institusi diharapkan dapat meningkatkan kemampuan generalisasi model pada kondisi data yang lebih beragam.
2. Seleksi fitur berbasis *Mutual Information* yang digunakan pada penelitian ini bersifat univariate dan tidak mempertimbangkan interaksi antar gen. Pendekatan seperti *Recursive Feature Elimination* (RFE), metode berbasis jaringan interaksi gen (*gene regulatory network*), atau teknik seleksi fitur berbasis *deep learning* seperti attention mechanism dapat dieksplorasi pada penelitian selanjutnya untuk mengidentifikasi biomarker yang lebih representatif dan bermakna secara biologis.
3. Penelitian selanjutnya dapat mengeksplorasi arsitektur *deep learning* yang lebih canggih, seperti *Convolutional Neural Network* (CNN) satu dimensi atau *Transformer-based* model, untuk menangkap pola sekuensial dan

dependensi jarak jauh antar fitur gen yang tidak dapat dimodelkan secara optimal oleh arsitektur MLP.

DAFTAR PUSTAKA

Akter, S., Adesola, R. O., & Basnet, S. (2025). *Machine learning* approach to identify significant genes and classify *cancer* types from RNA-seq data. *Global Medical Genetics*, *12*(4), 100079. https://doi.org/10,1016/j.gmg.2025.100079

Al Shaykh, 'Abdullah ibn Muhammad ibn 'Abd al-Rahman ibn Ishaq, & Abdul Ghoffar, M. (2006). *Tafsir Ibnu Katsir*. Pustaka Imam Asy-Syafi'i.

Alharbi, F., & Vakanski, A. (2023). *Machine learning* Methods for *Cancer* Classification Using *Gene expression* Data: A Review. *Bioengineering*, *10*(2), 173. https://doi.org/10,3390/bioengineering10020173

Alkawaz, A. N., Kanesan, J., Khairuddin, A. S. M., Badruddin, I. A., Kamangar, S., Hussien, M., Baig, M. A. A., & Ahammad, N. A. (2023). Training *Multilayer* Neural Network Based on Optimal Control Theory for Limited Computational Resources. *Mathematics*, *11*(3), 778. https://doi.org/10,3390/math11030778

Alzoubi, H., Alzubi, R., & Ramzan, N. (2023). *Deep learning* Framework for Complex Disease Risk Prediction Using Genomic Variations. *Sensors*, *23*(9), 4439. https://doi.org/10,3390/s23094439

Anggriani, R. A. H., Soeroso, N. N., Tarigan, S. P., Eyanoer, P. C., & Hidayat, H. (2022). Association of CYP2A6 Genetic Polymorphism and *Lung Cancer* in Female Never Smokers. *Molecular and Cellular Biomedical Sciences*, *6*(2), 63. https://doi.org/10,21705/mcbs.v6i2.232

Bargagli Stoffi, F. J., Cevolani, G., & Gnecco, G. (2022). Simple Models in Complex Worlds: Occam's Razor and Statistical Learning Theory. *Minds and Machines*, *32*(1), 13–42. https://doi.org/10,1007/s11023-022-09592-z

Bychowski, J., & Sobiczewski, W. (2023). Current perspectives of cardio-oncology: Epidemiology, adverse effects, pre-treatment screening and prevention strategies. *Cancer Medicine*, *12*(13), 14545–14555. https://doi.org/10,1002/cam4.5980

Deo, S. V. S., Sharma, J., & Kumar, S. (2022). GLOBOCAN 2020 Report on Global *Cancer* Burden: Challenges and Opportunities for Surgical Oncologists. *Annals of Surgical Oncology*, *29*(11), 6497–6500, https://doi.org/10,1245/s10434-022-12151-6

Du, K.-L., Leung, C.-S., Mow, W. H., & Swamy, M. N. S. (2022). *Perceptron*: Learning, Generalization, Model *Selection*, Fault Tolerance, and Role in the *Deep learning* Era. *Mathematics*, *10*(24), 4730, https://doi.org/10,3390/math10244730

Gu, Y. (2024). Prediction of *Lung Cancer* Stage Using Tumor *Gene expression* Data. *JouRNAl of Cancer Therapy*, *15*(08), 287–302. https://doi.org/10,4236/jct.2024.158028

Gupta, S., Gupta, M. K., Shabaz, M., & Sharma, A. (2022). *Deep learning* techniques for *cancer* classification using *Microarray gene expression* data. *Frontiers in Physiology*, *13*, 952709. https://doi.org/10,3389/fphys.2022.952709

Hanczar, B., Bourgeais, V., & Zehraoui, F. (2022). Assessment of *deep learning* and transfer learning for *cancer* prediction based on *gene expression* data. *BMC Bioinformatics*, *23*(1), 262. https://doi.org/10,1186/s12859-022-04807-7

Jiang, Z. (2024). MLP-Based *Lung Cancer* Prediction and *Feature* Importance Evaluation: *Proceedings of the 1st InteRNAtional Conference on Modern Logistics and Supply Chain Management*, 320–323. https://doi.org/10,5220/0013329900004558

Lau, S. C. M., Pan, Y., Velcheti, V., & Wong, K. K. (2022). Squamous cell *lung cancer*: Current landscape and future therapeutic options. *Cancer Cell*, *40*(11), 1279–1293. https://doi.org/10,1016/j.ccell.2022.09.018

Lima, D. V. C., Terrematte, P., Stransky, B., & Neto, A. D. D. (2025). *Explainable AI and Multiclassifiers for Staging Biomarker Discovery in Lung Squamous Cell Carcinoma*. *Cancer* Biology. https://doi.org/10,1101/2025.07.10,663226

Lin, L., & Bao, Y. (2024). Developent and validation of *machine learning* models for diagnosis and prognosis of *lung* adenocarcinoma, and immune infiltration analysis. *Scientific Reports*, *14*(1), 22081. https://doi.org/10,1038/s41598-024-73498-2

Luo, G., Zhang, Y., Rumgay, H., Morgan, E., Langselius, O., Vignat, J., Colombet, M., & Bray, F. (2025). Estimated worldwide variation and trends in incidence of *lung cancer* by histological subtype in 2022 and over time: A population-based study. *The Lancet Respiratory Medicine*, *13*(4), 348–363. https://doi.org/10,1016/S2213-2600(24)00428-4

Mohamed, T. I. A., Ezugwu, A. E., Fonou-Dombeu, J. V., Mohammed, M., Greeff, J., & Elbashir, M. K. (2023). A novel *feature Selection* algorithm for identifying hub genes in *lung cancer*. *Scientific Reports*, *13*(1), 21671. https://doi.org/10,1038/s41598-023-48953-1

Nassif, A. B., Abujabal, N. A., & Omar, A. A. (2025a). Classification of *lung cancer* severity using *gene expression* data based on *deep learning*. *BMC Medical Informatics and Decision Making*, *25*(1), 184. https://doi.org/10,1186/s12911-025-03011-w

Nassif, A. B., Abujabal, N. A., & Omar, A. A. (2025b). Classification of *lung cancer* severity using *gene expression* data based on *deep learning*. *BMC Medical Informatics and Decision Making*, *25*(1), 184. https://doi.org/10,1186/s12911-025-03011-w

noepinefrin. (2024). *TCGA - LUSC | Lung Cancer Gene expression Dataset*. Kaggle. https://www.kaggle.com/datasets/noepinefrin/tcga-lusc-*lung* -cell-squamous-carcinoma-gene-exp

Shah, S. N. A., Issar, K., & Parveen, R. (2026). A *hybrid feature extraction* framework combining PCA and *Mutual Information* for *gene expression* based *lung cancer* classification. *PLOS One*, *21*(2), e0342160, https://doi.org/10,1371/journal.pone.0342160

Sharma, R. (2022). Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of *lung cancer* in 2020 and 2050, *InteRNAtional JouRNAl of Clinical Oncology*, *27*(4), 665–675. https://doi.org/10,1007/s10147-021-02108-2

Shatravin, V., Shashev, D., & Shidlovskiy, S. (2022). Sigmoid Activation Implementation for Neural Networks Hardware Accelerators Based on Reconfigurable Computing Environments for Low-Power Intelligent Systems. *Applied Sciences*, *12*(10), 5216. https://doi.org/10,3390/app12105216

Sheet, S., Ghosh, R., & Ghosh, A. (2024). Recognition of *cancer* mediating genes using MLP-SDAE model. *Systems and Soft Computing*, *6*, 200079. https://doi.org/10,1016/j.sasc.2024.200079

Shen, J., Shi, J., Luo, J., Zhai, H., Liu, X., Wu, Z., Yan, C., & Luo, H. (2022). *Deep learning* approach for *cancer* subtype classification using high-dimensional *gene expression* data. *BMC Bioinformatics*, *23*(1), 430, https://doi.org/10,1186/s12859-022-04980-9

Tabassum, N., Kamal, M. A. S., Akhand, M. A. H., & Yamada, K. (2024a). *Cancer* Classification from *Gene expression* Using *Ensemble* Learning with an Influential *Feature Selection* Technique. *BioMedInformatics*, *4*(2), 1275–1288. https://doi.org/10,3390/biomedinformatics4020070

Tabassum, N., Kamal, M. A. S., Akhand, M. A. H., & Yamada, K. (2024b). *Cancer* Classification from *Gene expression* Using *Ensemble* Learning with an Influential *Feature Selection* Technique. *BioMedInformatics*, *4*(2), 1275–1288. https://doi.org/10,3390/biomedinformatics4020070

Wang, S., Zhang, H., Liu, Z., & Liu, Y. (2022). A Novel *Deep learning* Method to Predict *Lung Cancer* Long-Term Survival With Biological Knowledge Incorporated *Gene expression* Images and Clinical Data. *Frontiers in Genetics*, *13*, 800853. https://doi.org/10,3389/fgene.2022.800853

Yang, X., Zheng, Y., Xing, X., Sui, X., Jia, W., & Pan, H. (2022). Immune subtype identification and *Multi-layer Perceptron* classifier construction for breast *cancer*. *Frontiers in Oncology*, *12*, 943874. https://doi.org/10,3389/fonc.2022.943874

Yuvaraj, V., & Maheswari, D. (2023). *Lung cancer* classification based on enhanced *deep learning* using *gene expression* data. *Measurement: Sensors*, *30*, 100902. https://doi.org/10,1016/j.measen.2023.100902