

Terbit online pada laman : <http://teknosi.fti.unand.ac.id/>

# Jurnal Nasional Teknologi dan Sistem Informasi

| ISSN (Print) 2460-3465 | ISSN (Online) 2476-8812 |



Artikel Penelitian

## Perbandingan Metode Seleksi Fitur Chi-Square dan Information Gain untuk Peningkatan Interpretabilitas dan Optimasi Kinerja Model TabNet

Annisa Ratna Salsabilla <sup>a</sup>, Ramadhan Rakhmat Sani <sup>a\*</sup>, Ika Novita Dewi <sup>a</sup><sup>a</sup> Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro, Semarang 50131

### INFORMASI ARTIKEL

#### Sejarah Artikel:

Diterima Redaksi: 18 Oktober 2025

Revisi Akhir: 18 Desember 2025

Diterbitkan Online: 28 Desember 2025

### KATA KUNCI

Kanker Payudara,  
TabNet,  
Chi-Square,  
Information Gain,  
Optuna

### KORESPONDENSI

E-mail: ramadhan\_rs@dsn.dinus.ac.id \*

### A B S T R A C T

Kanker payudara merupakan salah satu masalah kesehatan global yang paling signifikan. Untuk membantu diagnosis dini, pendekatan machine learning menawarkan potensi untuk menganalisis data klinis secara akurat. Namun, model machine learning konvensional seringkali memiliki keterbatasan dalam memodelkan hubungan non-linear yang kompleks pada data medis, yang dapat membatasi akurasi prediktif. Penelitian ini menggunakan arsitektur deep learning karena kemampuannya dalam memodelkan hubungan non-linear yang kompleks. Secara spesifik, model TabNet dipilih karena dirancang untuk data tabular dan menawarkan interpretabilitas lebih baik. Penelitian ini menggunakan dataset publik Wisconsin Diagnostic Breast Cancer (WDBC) dengan 30 fitur dan distribusi kelas tidak seimbang, sehingga memerlukan seleksi fitur untuk menangani data berdimensi tinggi dan SMOTE-ENN untuk penyeimbangan kelas. Dua metode seleksi fitur, Chi-Square dan Information Gain, dibandingkan untuk menentukan pendekatan paling efektif. Untuk memastikan performa optimal, optimasi hiperparameter dilakukan menggunakan Optuna dan divalidasi dengan Stratified K-Fold Cross-Validation. Hasil eksperimen menunjukkan bahwa seleksi fitur dan optimasi secara signifikan meningkatkan kinerja. Model dasar dengan seleksi fitur Chi-Square menghasilkan akurasi 64.91%. Sedangkan model Chi-Square dengan optimasi Optuna meningkatkan akurasi sebesar 98.25%. Peningkatan ini juga lebih tinggi 3.51% dibandingkan dengan model yang dioptimalkan tanpa seleksi fitur dengan akurasi 94.74%. Dalam perbandingan akhir, kedua metode menunjukkan keunggulan berbeda yaitu Chi-Square (75% fitur) unggul dalam presisi 100% dan waktu komputasi lebih efisien, sementara Information Gain (75% fitur) menjadi satu-satunya yang mencapai recall 100%, yang krusial untuk meminimalkan false negative. Temuan ini menunjukkan bahwa pilihan metode optimal bersifat kontekstual. Information Gain unggul untuk sensitivitas diagnostik maksimal, sementara Chi-Square lebih unggul untuk keseimbangan performa dan efisiensi.

## 1. PENDAHULUAN

Kanker payudara merupakan jenis kanker yang paling umum terjadi pada perempuan di seluruh dunia. Menurut laporan *Global Cancer Observatory* terdapat lebih dari 2,3 juta kasus baru kanker payudara yang didiagnosis pada tahun 2022, menjadikannya kanker dengan insiden tertinggi kedua secara global setelah kanker paru-paru, serta tingkat mortalitas kanker payudara juga tinggi, dengan lebih dari 665.000 kematian per tahun yang menjadikan penyebab kematian kanker kedua tertinggi di kalangan perempuan [1]. Di tingkat regional, proyeksi

peningkatan jumlah kasus kanker payudara di Asia hingga 1,4 juta dan 0,5 juta kematian pada tahun 2050 menunjukkan perlunya strategi diagnostik dan prediksi yang lebih akurat untuk mengurangi beban kesehatan dan sosial yang ditimbulkan oleh penyakit ini [2]. Kanker payudara di Indonesia memiliki insidensi mencapai 42,1 kasus per 100.000 penduduk dan tingkat mortalitas mencapai 17 kematian per 100.000 penduduk, menjadikannya salah satu masalah kesehatan utama di negara ini [3]. Prevalensi yang tinggi ini sebagian besar disebabkan oleh deteksi yang terlambat, keterbatasan akses terhadap layanan kesehatan, dan kurangnya kesadaran akan pentingnya skrining dini [4].

Sejumlah penelitian telah mengeksplorasi berbagai model machine learning konvensional untuk klasifikasi kanker payudara menggunakan dataset publik *Wisconsin Diagnostic Breast Cancer* (WDBC) [5]. Dataset ini berisi fitur-fitur numerik yang dihitung dari karakteristik inti sel pada citra digital hasil prosedur *Fine Needle Aspiration* (FNA) [6]. Berbagai penelitian yang memanfaatkan dataset ini menunjukkan potensi signifikan dalam akurasi prediksi. Sebagai contoh, Chen et al. [7] melaporkan akurasi 97,4% menggunakan XGBoost, sementara Ara et al. [8] mencapai 96,5% dengan Random Forest dan Support Vector Machine (SVM). Penelitian lain oleh Albadr et al. [9] yang mengaplikasikan Fast Learning Network (FLN) juga memperoleh akurasi optimal hingga 98,83%. Lebih lanjut, penelitian terkini oleh Zhou et al. [10] berhasil mencapai akurasi 99,12% dengan model AdaBoost-Logistic, dan bahkan mengutip hasil 99,03% dari Abdur Rasool et al. menggunakan Polynomial SVM. Meskipun model-model machine learning ini menunjukkan kemampuan prediktif yang sangat baik, tantangan terkait interpretabilitas dan penanganan fitur yang kompleks secara inheren dalam data tabular masih menjadi perhatian.

Untuk menjawab keterbatasan tersebut, penelitian ini mengusulkan penggunaan TabNet, sebuah model deep learning yang dirancang khusus untuk menangani data tabular dengan mekanisme perhatian yang memungkinkan pemilihan fitur secara adaptif, sehingga lebih transparan dan dapat diinterpretasi dibandingkan model deep learning konvensional [11]. Dengan pendekatan ini, TabNet tidak hanya fokus pada akurasi prediksi, tetapi juga pada interpretabilitas, menjadikannya lebih sesuai untuk aplikasi medis seperti klasifikasi kanker payudara [12]. Namun, sebagai sebuah arsitektur deep learning, efektivitas TabNet sangat bergantung pada kualitas data input. Model ini tidak secara inheren dirancang untuk menyaring fitur yang tidak relevan atau menangani data yang sangat tidak seimbang, yang merupakan tantangan umum pada dataset medis. Oleh karena itu, untuk memaksimalkan potensi TabNet, diperlukan sebuah alur pra-pemrosesan data yang cermat.

Penerapan seleksi fitur menjadi langkah krusial untuk mengidentifikasi dan mempertahankan hanya fitur-fitur yang paling relevan secara informatif. Pendekatan ini bertujuan untuk meningkatkan akurasi, mengurangi risiko *overfitting*, dan mempercepat waktu pelatihan model. Penggunaan model TabNet dan pendekatan feature selection tersebut belum banyak diterapkan untuk klasifikasi kanker payudara, sehingga menyisakan celah penting untuk dieksplorasi. Secara spesifik, penelitian ini akan menerapkan dan membandingkan dua pendekatan seleksi fitur dari kategori filter yang relevan untuk

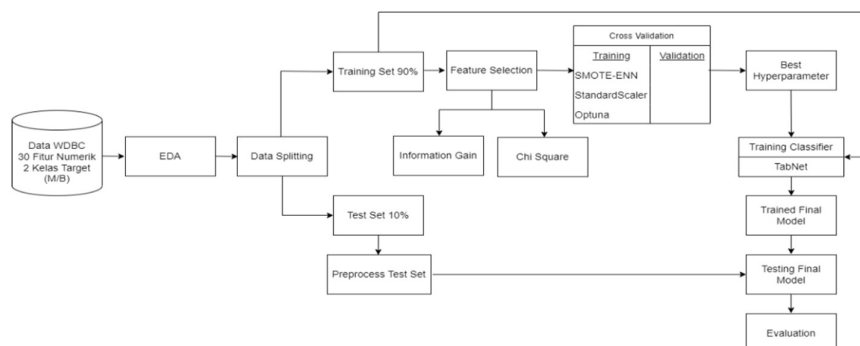
ugas klasifikasi yaitu Information Gain dan Chi-Square. Information Gain, yang berakar dari teori informasi, akan digunakan untuk mengukur nilai informatif setiap fitur dengan mengevaluasi besarnya pengurangan ketidakpastian pada kelas target yang dihasilkan oleh fitur [13]. Sementara itu, Chi-Square akan digunakan untuk menguji ketergantungan statistik antara setiap fitur dan kelas target [14].

Untuk mengatasi bias yang mungkin timbul dari data yang tidak seimbang, penelitian ini menerapkan Synthetic Minority Over-sampling Technique-Edited Nearest Neighbor (SMOTE-ENN), sebuah teknik resampling hibrida yang menggabungkan dua metode sekuensial yaitu Synthetic Minority Over-sampling Technique (SMOTE) sebagai metode *oversampling* dan Edited Nearest Neighbors (ENN) sebagai metode *undersampling* [15]. Teknik ini terbukti efektif karena SMOTE terlebih dahulu menyintesis sampel baru untuk kelas minoritas guna mengatasi ketidakseimbangan data, yang kemudian diikuti oleh ENN yang membersihkan data dengan menghapus sampel yang dianggap bising (*noisy*) untuk menciptakan batas keputusan yang lebih jelas [16]. Selanjutnya, untuk memastikan model TabNet beroperasi pada performa puncaknya, proses optimasi hiperparameter dilakukan menggunakan Optuna, sebuah kerangka kerja yang terbukti mampu menyederhanakan proses pencarian parameter secara sistematis pada model kompleks seperti TabNet untuk mencapai performa yang optimal [17].

Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model klasifikasi biner kanker payudara berbasis TabNet yang dioptimalkan melalui alur kerja komprehensif mencakup seleksi fitur, penyeimbangan data, dan optimasi hiperparameter. Secara spesifik, model akan dilatih untuk membedakan antara dua kelas target yaitu jinak (*Benign*) dan ganas (*Malignant*). Model TabNet akan dilatih menggunakan subset fitur hasil seleksi dari masing-masing pendekatan ini, dan kinerjanya akan dievaluasi berdasarkan akurasi, presisi, recall, dan f1-score. Kontribusi utama dari penelitian ini adalah menyajikan bukti empiris mengenai efektivitas komparatif antara Information Gain dan Chi-Square sebagai teknik pra-pemrosesan untuk model TabNet, yang tidak hanya bertujuan meningkatkan performa prediktif tetapi juga mendukung transparansi model.

## 2. METODE

Metode penelitian yang digunakan diilustrasikan secara visual pada diagram alir Gambar 1. Alur kerja dimulai dengan tahap preparasi dan analisis data eksploratif (Exploratory Data Analysis – EDA). Setelah itu, data dibagi menjadi data latih



Gambar 1. Diagram Alur Penelitian

(90%) dan data uji (10%). Selanjutnya, serangkaian proses diterapkan secara eksklusif pada data latih, yang mencakup penskalaan fitur menggunakan StandardScaler, penyeimbangan kelas dengan SMOTE-ENN, dan perbandingan efektivitas metode seleksi fitur melalui dua skenario utama yaitu Chi-Square dan Information Gain. Proses optimasi hiperparameter tuning optuna model divalidasi menggunakan teknik Stratified 5-Fold Cross-Validation. Seluruh implementasi teknis dibangun di atas platform Google Colaboratory menggunakan bahasa pemrograman Python.

## 2.1 Dataset

Penelitian ini menggunakan dataset public *Wisconsin Diagnostic Breast Cancer* (WDBC) yang diperoleh dari UCI Machine Learning Repository [6]. Dataset ini terdiri dari 569 sampel data yang terbagi ke dalam dua kelas, yaitu 357 sampel diklasifikasikan sebagai *Benign* (jinak) dan 212 sampel sebagai *Malignant* (ganas). Setiap sampel direpresentasikan oleh 30 fitur numerik yang dihitung dari citra digital hasil prosedur *Fine Needle Aspiration* (FNA) dan menggambarkan karakteristik inti sel. Fitur-fitur ini secara struktural dikelompokkan ke dalam tiga kategori perhitungan statistik untuk sepuluh pengukuran inti sel yang berbeda yaitu nilai rata-rata (*mean*), galat baku (*standard error*), dan nilai terburuk (*worst*). Kesepuluh pengukuran inti sel tersebut meliputi radius, tekstur, perimeter, area, kehalusan (*smoothness*), kepadatan (*compactness*), cekungan (*concavity*), titik cekung (*concave points*), simetri, dan dimensi fraktal [6].

## 2.2 Pra-pemrosesan Data

Sebelum dilakukan seleksi fitur dan pemodelan, data mentah melalui beberapa tahap pra-pemrosesan. Tahap ini diawali dengan pembersihan data, yaitu menghapus kolom 'ID' yang tidak memiliki nilai prediktif. Selanjutnya, variabel target yang bersifat kategorikal diubah menjadi representasi numerik biner melalui proses *label encoding* ('Malignant' menjadi 1 dan 'Benign' menjadi 0).

Seluruh 30 fitur numerik dalam dataset distandarisasi menggunakan StandardScaler. Teknik ini mentransformasikan setiap fitur sehingga memiliki rata-rata 0 dan deviasi standar 1 [7]. Penskalaan ini penting untuk memastikan bahwa semua fitur berada pada skala yang sebanding dan dapat diproses secara adil oleh model [10].

Terakhir, untuk menangani masalah ketidakseimbangan kelas pada dataset WDBC yang dapat menyebabkan model menjadi bias terhadap kelas mayoritas, teknik resampling hibrida SMOTE-ENN diterapkan secara eksklusif pada data training untuk menyeimbangkan distribusi sampel antar kelas. Pada tahap SMOTE, jumlah sampel pada kelas minoritas ditambah dengan cara menciptakan sampel sintetis baru. Sampel-sampel ini dibuat di antara sampel minoritas yang ada dan tetangga terdekatnya, sehingga menghasilkan data minoritas yang lebih beragam. Sedangkan tahap undersampling ENN dilakukan setelah tahap oversampling SMOTE. Tahap ini membersihkan data dengan menghapus sampel dari kedua kelas yang dianggap sebagai *noise* atau salah klasifikasi oleh tetangga terdekatnya. Kombinasi ini menghasilkan dataset latih yang tidak hanya

seimbang tetapi juga memiliki batas keputusan antar kelas yang lebih jelas, yang diterapkan secara eksklusif pada data training selama proses validasi silang.

## 2.3 Seleksi Fitur

Tahap seleksi fitur dilakukan untuk mengidentifikasi dan memilih subset fitur yang paling relevan terhadap variabel target.

### 2.3.1 Chi-Square

Chi-Square adalah uji statistik non-parametrik yang digunakan untuk mengevaluasi ketergantungan antara fitur dengan kelas target [18]. Nilai  $\chi^2$  yang tinggi mengindikasikan bahwa suatu fitur tidak independen dari kelas, sehingga dianggap relevan untuk klasifikasi [14]. Uji Chi-Square secara fundamental memerlukan input fitur yang bersifat non-negatif [19]. Oleh karena itu, meskipun StandardScaler digunakan untuk pra-pemrosesan umum, secara khusus untuk tahap seleksi fitur ini diterapkan MinMaxScaler. Transformasi ini memastikan semua nilai fitur berada dalam rentang sebelum perhitungan statistik dilakukan. Rumus statistik Chi-Square didefinisikan pada rumus (1).

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

Dimana :

- $\chi^2$  = Nilai statistik Chi-Square.
- $O_i$  = Frekuensi yang diamati pada kategori ke- $i$ .
- $E_i$  = Frekuensi yang diharapkan pada kategori ke- $i$
- $n$  = Jumlah total kategori.

### 2.3.2 Information Gain

Information Gain (IG) adalah metode seleksi fitur yang mengukur tingkat dependensi antara dua variabel berdasarkan teori informasi, IG juga mampu menangkap hubungan linear maupun non-linear, di mana nilai IG yang tinggi menunjukkan tingkat ketergantungan yang kuat antara fitur dan kelas target [13]. Perhitungan Mutual Information antara fitur  $X$  dan target  $Y$  diberikan oleh rumus (2).

$$I(X; Y) = H(X) - H(X|Y) \quad (2)$$

Di mana :

- $I(X; Y)$  = Mutual Information antara fitur  $X$  dan target  $Y$ .
- $H(X)$  = Entropi dari target  $Y$ .
- $H(X|Y)$  = Entropi kondisional dari  $Y$  setelah  $X$  diobservasi.

## 2.4 Metode Klasifikasi

Model klasifikasi utama yang diimplementasikan dalam penelitian ini adalah TabNet, sebuah arsitektur *deep learning* yang dirancang secara spesifik untuk data tabular [20]. Model ini diimplementasikan menggunakan kelas TabNetClassifier dari pustaka pytorch-tabnet. Arsitektur TabNet memanfaatkan mekanisme perhatian sekuensial (*sequential attention*) untuk memilih subset fitur paling relevan pada setiap langkah keputusan (*decision step*) [11]. Pendekatan ini memungkinkan model untuk melakukan seleksi fitur *instance-*

wise, yaitu memilih fitur yang berbeda untuk setiap sampel data, sehingga meningkatkan interpretabilitas dan efisiensi pemodelan [12].

Berdasarkan arsitektur yang diusulkan oleh Arik dan Pfister [21], alur kerja internal TabNet untuk memproses setiap sampel data dapat diuraikan menjadi beberapa tahapan utama sebagai berikut:

1. Pemrosesan Fitur Awal: Pertama, seluruh fitur masukan dilewatkan melalui sebuah lapisan *Batch Normalization*. Fitur yang telah dinormalisasi ini kemudian diproses oleh sebuah blok *Feature Transformer* yang lapisannya digunakan bersama di seluruh langkah keputusan. Tujuannya adalah untuk mengekstraksi representasi fitur awal yang kaya.
2. Langkah Keputusan Sekuensial: Model kemudian memasuki serangkaian langkah keputusan yang berurutan sesuai hiperparameter  $n\_steps$ . Pada setiap langkah, terjadi dua proses kunci. Pertama seleksi fitur oleh *Attentive Transformer* yang menerima informasi dari langkah sebelumnya untuk menghasilkan sebuah *attention mask*. *Mask* ini, melalui fungsi aktivasi *sparsemax*, secara dinamis memilih subset fitur yang paling relevan untuk dianalisis pada langkah saat ini. Proses ini memastikan bahwa kapasitas pemodelan difokuskan hanya pada fitur-fitur yang paling informatif. Kedua, pemrosesan fitur oleh *Feature Transformer* dimana fitur yang telah dipilih oleh *attention mask* kemudian diproses oleh blok *Feature Transformer* yang spesifik untuk langkah tersebut (*step-dependent*). Blok ini bertugas untuk mengekstraksi informasi dari subset fitur yang relevan untuk berkontribusi pada keputusan akhir.
3. Agregasi Keputusan: Output dari setiap blok *Feature Transformer* di setiap langkah keputusan kemudian diagregasi atau digabungkan. Keputusan klasifikasi akhir tidak hanya didasarkan pada satu langkah, melainkan pada akumulasi informasi dari seluruh rangkaian langkah keputusan.
4. Manajemen Atensi: Mekanisme perhatian TabNet juga dirancang untuk melacak seberapa sering setiap fitur telah digunakan pada langkah-langkah sebelumnya. Hal ini mendorong model untuk mengeksplorasi fitur-fitur yang berbeda di setiap langkah, menghasilkan pemodelan yang lebih holistik dan mencegah ketergantungan berlebihan pada beberapa fitur dominan.

Arsitektur multi-langkah ini memungkinkan TabNet untuk membangun model yang sangat non-linear dari data tabular, sambil tetap menyediakan interpretabilitas melalui visualisasi *attention mask* yang menunjukkan fitur mana yang paling penting pada setiap tahap pengambilan keputusan.

## 2.5 Optimasi Hiperparameter Model

Tahap optimasi bertujuan untuk menemukan konfigurasi hiperparameter terbaik untuk model TabNet secara sistematis. Proses ini dilakukan secara otomatis menggunakan kerangka kerja Optuna sebanyak 30 *trial*. Untuk setiap *trial*, kinerja dari satu set hiperparameter dievaluasi menggunakan metode Stratified K-Fold Cross-Validation ( $k=5$ ). Penggunaan Stratified K-Fold memastikan bahwa setiap *fold* data memiliki proporsi kelas target yang representatif. Tujuan dari optimasi ini

adalah untuk memaksimalkan akurasi rata-rata dari kelima *fold* validasi tersebut.

Untuk menjaga validitas dari setiap evaluasi dan mencegah kebocoran data (*data leakage*), sebuah *pipeline* pra-pemrosesan diintegrasikan di dalam setiap iterasi *cross-validation*. Artinya, untuk setiap *fold*, proses *fitting* dari StandardScaler dan penerapan *resampling* dengan SMOTE-ENN hanya dilakukan pada data latih-fold. Data validasi-fold kemudian hanya ditransformasi menggunakan parameter scaler yang telah di-*fit* dari data latih-fold tersebut. Ruang pencarian hiperparameter yang dieksplorasi mencakup parameter kunci dari arsitektur TabNet ( $n\_d, n\_a, n\_steps, \gamma, \lambda, \text{lambda\_sparse}$ ) dan parameter optimizer Adam ( $\text{lr}, \text{betas}$ ).

## 2.6 Evaluasi Model

Kinerja model final dievaluasi pada *Test Set* menggunakan metrik-metrik standar yang diturunkan dari Confusion Matrix. Komponen dari *confusion matrix* meliputi *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [22]. Berdasarkan komponen tersebut, empat metrik utama dihitung untuk menilai performa model secara komprehensif.

Akurasi mengukur proporsi total prediksi yang benar dari keseluruhan sampel. Meskipun umum digunakan, akurasi dapat menyesatkan pada dataset yang tidak seimbang [22].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Presisi dikenal sebagai *Positive Predictive Value* (PPV). Mengukur proporsi prediksi positif yang benar dari total prediksi positif yang dibuat oleh model [22].

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

Recall juga dikenal sebagai Sensitivitas (*Sensitivity*) atau *True Positive Rate* (TPR). Mengukur proporsi kasus positif aktual yang berhasil diidentifikasi oleh model [22].

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

F1-Score merupakan rata-rata harmonik (*harmonic mean*) dari Presisi dan Recall. F1-Score memberikan skor tunggal yang menyeimbangkan kedua metrik tersebut dan sangat berguna untuk mengevaluasi model pada dataset yang tidak seimbang [22].

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

## 3. HASIL

### 3.1 Hasil Eksplorasi Dataset

Seluruh Dataset WDBC yang digunakan dalam penelitian ini terdiri dari 569 sampel dan 32 kolom. Setelah dilakukan penghapusan kolom 'ID' yang tidak relevan pada tahap pra-pemrosesan, dataset final yang digunakan untuk eksperimen

terdiri dari 569 sampel, 30 fitur prediktif, dan 1 kolom target. Distribusi kelas pada dataset ini tidak seimbang, dengan 357 sampel (62.7%) termasuk dalam kelas Benign (jinak) dan 212 sampel (37.3%) dalam kelas Malignant (ganas).

Fitur-fitur dalam dataset ini merupakan pengukuran kuantitatif dari karakteristik inti sel. Untuk memberikan gambaran yang lebih jelas, Tabel 1 menyajikan ringkasan dari 10 pengukuran inti sel, beserta deskripsi dan statistik deskriptifnya.

Tabel 1. Deskripsi dan Statistik Fitur Inti pada Dataset WDBC

Pengukuran Inti	Deskripsi Singkat	Mean	Std Dev	Rentang Nilai (Min - Max)
Radius	Jarak rata-rata dari pusat ke keliling	14.127	3.524	6.981 - 28.11
Texture	Deviasi standar dari nilai skala abu-abu	19.290	4.301	9.71 - 39.28
Perimeter	Keliling dari inti sel	91.969	24.299	43.79 - 188.5
Area	Luas dari inti sel	654.889	351.914	143.5 - 2501.0
Smoothness	Variasi lokal dalam panjang radius	0.096	0.014	0.053 - 0.163
Compactness	$\text{Perimeter}^2 / \text{Area} - 1.0$	0.104	0.053	0.019 - 0.345
Concavity	Tingkat keparahan kontur cekung	0.089	0.080	0.000 - 0.427
Concave Points	Jumlah bagian kontur yang cekung	0.049	0.039	0.000 - 0.201
Symmetry	Simetri dari inti sel	0.181	0.027	0.106 - 0.304
Fractal Dimension	"Perkiraan garis pantai" - 1	0.063	0.007	0.050 - 0.097

Seperti yang terlihat pada Tabel 1, terdapat variasi skala yang sangat signifikan antar fitur. Sebagai contoh, fitur `area_mean` memiliki nilai rata-rata lebih dari 650, sedangkan `smoothness_mean` memiliki nilai rata-rata kurang dari

0.1. Variasi ini memvalidasi pentingnya penerapan penskalaan fitur pada tahap pra-pemrosesan untuk memastikan setiap fitur memberikan kontribusi yang adil dalam proses pemodelan.

### 3.2 Hasil Peringkat Fitur Berdasarkan Chi-Square

Pada skenario pertama, seleksi fitur dilakukan menggunakan uji statistik Chi-Square (CS) untuk mengidentifikasi fitur-fitur yang paling berpengaruh. Dari total 30 fitur, hasil perhitungan menunjukkan bahwa tiga fitur dengan tingkat ketergantungan tertinggi terhadap kelas target adalah `concave_points1` (skor 47.70), `concavity1` (skor 41.84), dan `concave_points3` (skor 41.58).

Untuk mengevaluasi dampak dari jumlah fitur terhadap kinerja model secara sistematis, penelitian ini melakukan pengujian bertingkat dengan menggunakan persentase *cut-off* yang berbeda berdasarkan peringkat skor Chi-Square. Pendekatan ini bertujuan untuk menemukan jumlah fitur optimal yang dapat memaksimalkan performa prediktif. Tiga level *cut-off* yang digunakan dalam eksperimen adalah 25% (7 fitur teratas), 50% (15 fitur teratas), dan 75% (22 fitur teratas) dari total fitur. Tabel 2 menyajikan peringkat lengkap dari 30 fitur beserta skor Chi-Square-nya, yang menjadi dasar untuk penentuan subset fitur pada setiap level *cut-off*.

Tabel 2. Hasil Peringkat Fitur Berdasarkan Skor Chi-Square

No	Fitur	Skor CS
0	<code>concave_points1</code>	47.704.702
1	<code>concavity1</code>	41.841.923
2	<code>concave_points3</code>	41.577.207
3	<code>area3</code>	31.843.266
4	<code>perimeter3</code>	31.299.567
5	<code>radius3</code>	30.855.405
6	<code>concavity3</code>	27.789.702
7	<code>area1</code>	26.475.035
8	<code>perimeter1</code>	23.856.547
9	<code>radius1</code>	22.323.021
10	<code>compactness1</code>	19.186.956
11	<code>compactness3</code>	18.876.998
12	<code>area2</code>	18.430.961
13	<code>radius2</code>	16.423.247
14	<code>perimeter2</code>	15.406.083
15	<code>texture3</code>	8.061.342
16	<code>compactness2</code>	6.100.187
17	<code>texture1</code>	5.895.706
18	<code>concave_points2</code>	5.466.587
19	<code>smoothness3</code>	5.394.156
20	<code>symmetry3</code>	4.911.462
21	<code>fractal_dimension3</code>	4.200.376
22	<code>smoothness1</code>	4.002.087
23	<code>symmetry1</code>	2.695.648
24	<code>concavity2</code>	2.363.480
25	<code>fractal_dimension2</code>	0.355056
26	<code>smoothness2</code>	0.068104
27	<code>texture2</code>	0.000572
28	<code>symmetry2</code>	0.000476
29	<code>fractal_dimension1</code>	0.000206

### 3.3 Hasil Peringkat Fitur Berdasarkan Information Gain



Pada skenario kedua, seleksi fitur dilakukan menggunakan metode Information Gain (IG) untuk mengukur kontribusi informatif setiap fitur terhadap kelas target. Hasil perhitungan menunjukkan bahwa tiga fitur dengan skor IG tertinggi secara berurutan adalah perimeter3 (skor 0.455), area3 (skor 0.450), dan radius3 (skor 0.448).

Sama seperti pada skenario Chi-Square, untuk mengevaluasi dampak dari jumlah fitur terhadap kinerja model, dilakukan pengujian bertingkat. Pendekatan ini menggunakan persentase cut-off yang sama berdasarkan peringkat skor Information Gain. Tiga level cut-off yang digunakan dalam eksperimen adalah 25% (7 fitur teratas), 50% (15 fitur teratas), dan 75% (22 fitur teratas) dari total fitur. Tabel 3 menyajikan peringkat lengkap dari 30 fitur beserta skor Information Gain-nya.

Tabel 3. Hasil Peringkat Fitur Berdasarkan Skor Information Gain

No	Fitur	Skor IG
0	perimeter3	0.455044
1	area3	0.450911
2	radius3	0.448345
3	concave_points3	0.440064
4	concave_points1	0.436302
5	perimeter1	0.405262
6	concavity1	0.375536
7	radius1	0.363380
8	area1	0.351654
9	area2	0.338593
10	concavity3	0.307412
11	perimeter2	0.257667
12	radius2	0.237525
13	compactness3	0.220550
14	compactness1	0.205479
15	concave_points2	0.138218
16	texture3	0.130867
17	concavity2	0.111006
18	symmetry3	0.096550
19	smoothness3	0.091240
20	texture1	0.087187
21	smoothness1	0.083635
22	fractal_dimension3	0.076019
23	symmetry1	0.068308
24	compactness2	0.059189
25	fractal_dimension2	0.052337
26	symmetry2	0.018842
27	smoothness2	0.017293
28	fractal_dimension1	0.014088
29	texture2	0.000000

### 3.4 Kinerja Model Tanpa Seleksi Fitur

Sebelum mengevaluasi model dengan fitur yang telah diseleksi, penting untuk menetapkan kinerja dasar (*baseline performance*) dari model TabNet yang menggunakan seluruh 30 fitur. Tabel 4 menyajikan hasil evaluasi pada *Test Set* untuk dua kondisi yaitu model dengan hiperparameter *default* (tanpa optimasi) dan model yang telah dioptimalkan menggunakan Optuna.

Tabel 4. Kinerja Model Tanpa Seleksi Fitur

Kondisi Optimasi	Akurasi	Recall	Presisi	Waktu Komputasi
Tanpa Optimasi	82.46%	71.43%	78.95%	2.3662 s
Optimasi Optuna	94.74%	95.24%	90.91%	66.3101 s

Hasil pada Tabel 4 menunjukkan perbandingan kinerja antara kedua kondisi. Pada kondisi tanpa optimasi, model mencatat akurasi sebesar 82.46%, recall 71.43%, dan presisi 78.95% dengan waktu komputasi 2.37 detik. Setelah optimasi hiperparameter, tercatat adanya peningkatan pada semua metrik, di mana akurasi mencapai 94.74%, recall 95.24%, dan presisi 90.91%, dengan waktu komputasi 66.31 detik.

### 3.5 Kinerja Model dengan Seleksi Fitur Chi-Square

Implementasi seleksi fitur dengan Chi-Square, yang dikombinasikan dengan optimasi hiperparameter, menghasilkan beberapa model dengan kinerja yang bervariasi tergantung pada jumlah fitur yang digunakan. Pada kondisi tanpa optimasi, performa model cenderung rendah. Sebaliknya, setelah optimasi, semua level *cut-off* menunjukkan peningkatan performa yang substansial. Rincian lengkap hasil evaluasi disajikan pada Tabel 5.

Tabel 5. Kinerja Model dengan Seleksi Fitur Chi-Square

Kondisi Optimasi	Jumlah Fitur	Akurasi	Recall	Presisi	Waktu Komputasi
Tanpa Optimasi	25% (7 Fitur)	29.82%	76.19%	31.37%	8.2657 s
	50% (15 Fitur)	61.40%	0%	0%	7.1095 s
	75% (22 Fitur)	64.91%	4.76%	100%	4.7770 s
Optimasi Optuna	25% (7 Fitur)	100%	100%	100%	48.7241 s
	50% (15 Fitur)	96.49%	95.24%	95.24%	30.5398 s
	75% (22 Fitur)	<b>98.25%</b>	<b>95.24%</b>	<b>100%</b>	<b>29.2958 s</b>

Pada kondisi tanpa optimasi, hasil kinerja tercatat bervariasi dengan model dengan 25% fitur (7 fitur) memiliki akurasi 29.82% dan recall 76.19%, sedangkan model dengan 50% dan 75% fitur menghasilkan akurasi masing-masing 61.40% dan 64.91% dengan nilai recall yang sangat rendah. Sebaliknya, pada kondisi dengan optimasi, terjadi peningkatan performa yang signifikan. Model dengan 25% fitur mencapai skor 100% untuk semua metrik (akurasi, recall, dan presisi). Model dengan 50% fitur (15 fitur) memiliki akurasi 96.49%, recall 95.24%, dan presisi 95.24%. Selanjutnya, model dengan 75% fitur (22 fitur) menghasilkan akurasi 98.25%, recall 95.24%, dan presisi 100%.

### 3.6 Kinerja Model dengan Seleksi Fitur Information Gain

Serupa dengan skenario sebelumnya, penerapan seleksi fitur Information Gain menunjukkan bahwa optimasi hiperparameter secara signifikan meningkatkan kinerja model. Tanpa optimasi, akurasi tertinggi yang dicapai adalah 63.16%, sedangkan setelah optimasi, akurasi meningkat hingga 98.25%. Berikut adalah hasil

lengkap pemodelan dengan seleksi fitur Information Gain pada Tabel 6.

Tabel 6. Kinerja Model dengan Seleksi Fitur Information Gain

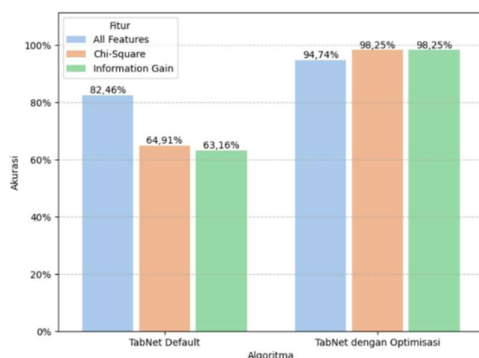
Kondisi Optimasi	Jumlah Fitur	Akurasi	Recall	Presisi	Waktu Komputasi
Tanpa Optimasi	25% (7 Fitur)	35.09%	95.24%	35.71%	2.7349 s
	50% (15 Fitur)	63.16%	0%	0%	2.4928 s
	75% (22 Fitur)	63.16%	0%	0%	4.1457 s
Optimasi Optuna	25% (7 Fitur)	94.74%	90.48%	95%	88.1092 s
	50% (15 Fitur)	<b>98.25%</b>	95.24%	<b>100%</b>	<b>35.9958 s</b>
	75% (22 Fitur)	<b>98.25%</b>	<b>100%</b>	95.45%	80.2105 s

Pada kondisi tanpa optimasi, model dengan 25% fitur (7 fitur) menghasilkan akurasi 35.09% dengan recall 95.24%, namun pada level 50% dan 75% fitur, model mencatat akurasi 63.16% dengan nilai recall 0%. Setelah optimasi, model dengan 25% fitur memiliki akurasi 94.74%, recall 90.48%, dan presisi 95%. Model dengan 50% fitur (15 fitur) mencapai akurasi 98.25%, recall 95.24%, dan presisi 100%. Terakhir, model dengan 75% fitur (22 fitur) juga memperoleh akurasi 98.25%, namun dengan recall 100% dan presisi 95.45%.

## 4. PEMBAHASAN

### 4.1 Pengaruh Optimasi Hiperparameter Optuna

Temuan pertama dari hasil eksperimen adalah peran penting dari optimasi hiperparameter tuning optuna dalam menentukan kinerja model TabNet. Pengaruh ini divisualisasikan secara jelas pada Gambar 2, yang membandingkan akurasi model pada kondisi *default* (tanpa optimasi) dengan kondisi yang telah dioptimalkan menggunakan Optuna.



Gambar 2. Perbandingan Akurasi Model Sebelum dan Sesudah Optimasi

Seperti yang diilustrasikan pada Gambar 2, proses optimasi memberikan peningkatan kinerja yang sangat besar di semua skenario. Pada model dasar tanpa seleksi fitur (*All Features*), akurasi meningkat dari 82.46% menjadi 94.74%. Dampak yang lebih terlihat yaitu pada skenario yang melibatkan seleksi fitur. Tanpa optimasi, model yang dilatih pada subset fitur yang lebih kecil bahkan menunjukkan kinerja yang lebih buruk daripada model *default* dengan akurasi 64.91% untuk Chi-Square dan 63.16% untuk Information Gain. Namun, setelah melalui proses optimasi, akurasi untuk kedua skenario seleksi fitur meningkat hingga 98.25%.

Analisis ini menegaskan bahwa pada dataset WDBC, proses optimasi hiperparameter optuna bukan hanya sekadar langkah penyempurnaan, melainkan menjadi faktor penentu keberhasilan untuk membuka potensi penuh dari arsitektur TabNet, terutama ketika dikombinasikan dengan seleksi fitur. Temuan ini mengindikasikan bahwa seleksi fitur hanya dapat memberikan manfaat jika diimplementasikan bersamaan dengan konfigurasi model yang tepat yaitu melalui proses *tuning* yang sistematis.

### 4.2 Efektifitas Seleksi Fitur

Setelah menetapkan pentingnya optimasi hiperparameter, analisis selanjutnya difokuskan pada efektivitas penerapan seleksi fitur. Hasil eksperimen menunjukkan bahwa penerapan seleksi fitur secara umum mampu meningkatkan kinerja prediktif model TabNet yang telah dioptimalkan. Model yang menggunakan seluruh 30 fitur mencapai akurasi 94.74%, sedangkan model yang menggunakan subset fitur yang lebih kecil berhasil mencapai akurasi hingga 98.25%. Peningkatan ini memvalidasi hipotesis bahwa penghilangan fitur yang kurang relevan dapat membantu model fokus pada sinyal yang paling informatif dan meningkatkan kemampuan generalisasi model. Gambar 3 menyajikan perbandingan akurasi secara visual antara skenario Chi-Square dan Information Gain pada berbagai tingkat *cut-off*.

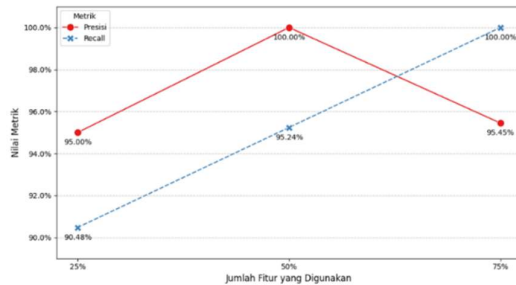


Gambar 3. Perbandingan Akurasi Model dengan Berbagai Skenario Seleksi Fitur

Berdasarkan Gambar 3, terlihat bahwa kedua metode seleksi fitur mampu mengidentifikasi subset fitur yang sangat prediktif, dengan keduanya berhasil mencapai akurasi tertinggi yang setara, yaitu 98.25%. Skenario Chi-Square menunjukkan performa yang sangat kuat pada *cut-off* 25% dengan akurasi sempurna 100%. Namun hasil sempurna seperti ini perlu diinterpretasikan dengan hati-hati. Dalam konteks eksperimen dengan *Test Set* yang berukuran kecil, skor 100% dapat menjadi indikasi adanya *overfitting*, di mana model menjadi terlalu spesifik terhadap data uji dan berpotensi tidak dapat digeneralisasi dengan baik pada data baru.

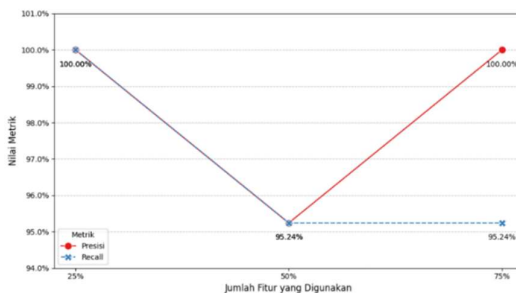
Oleh karena itu, kinerja Chi-Square yang lebih robust adalah model dengan *cut-off* 75% (22 fitur), yang mencapai akurasi tinggi sebesar 98.25%. Di sisi lain, Information Gain juga menunjukkan kinerja yang sangat kompetitif, mencapai akurasi tertinggi yang sama yaitu 98.25% pada level *cut-off* 50% dan 75%. Hal ini mengindikasikan bahwa kedua metode seleksi fitur mampu mengidentifikasi subset fitur yang sangat prediktif.

Untuk memahami perilaku model secara lebih mendalam, analisis Presisi dan Recall dilakukan untuk setiap skenario. Gambar 4 mengilustrasikan Presisi dan Recall untuk skenario Information Gain.



Gambar 4. Perbandingan Presisi dan Recall pada Skenario Information Gain

Seperti yang ditunjukkan pada Gambar 4, skenario Information Gain menunjukkan adanya perilaku *trade-off* yang jelas. Pada *cut-off* 50%, model ini mencapai Presisi 100% dengan Recall 95.24%. Namun, seiring dengan penambahan jumlah fitur ke 75%, Recall meningkat hingga 100%, tetapi dengan mengorbankan Presisi yang menurun menjadi 95.45%. Hal ini mengindikasikan bahwa dengan lebih banyak fitur, model IG menjadi lebih sensitif dalam mengidentifikasi semua kasus positif, namun dengan risiko menghasilkan lebih banyak prediksi positif yang keliru. Selanjutnya, Gambar 5 menyajikan analisis yang sama untuk skenario Chi-Square.



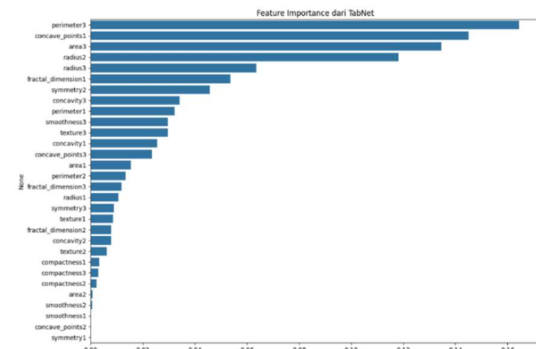
Gambar 5. Perbandingan Stabilitas Presisi dan Recall pada Skenario Chi-Square

Berbeda dengan Information Gain, Gambar 5 menunjukkan bahwa skenario Chi-Square memiliki perilaku yang lebih stabil. Ketika jumlah fitur ditingkatkan, model Chi-Square tidak menunjukkan adanya *trade-off* yang signifikan. Sebaliknya, model ini mampu mencapai Presisi sempurna (100%) pada *cut-off* 75% fitur, dengan nilai Recall yang tetap tinggi (95.24%). Stabilitas ini menunjukkan bahwa fitur-fitur yang dipilih oleh Chi-Square menghasilkan batas keputusan yang lebih konsisten.

Analisis akhir menunjukkan bahwa pilihan metode seleksi fitur yang terbaik bergantung pada prioritas tujuan. Jika tujuannya adalah keseimbangan performa dan efisiensi, maka Chi-Square dengan 75% fitur adalah konfigurasi yang unggul. Model ini mencapai performa yang identik dengan model Information Gain (50% fitur) yang juga mencapai akurasi 98.25% dan presisi 100% namun waktu komputasi yang lebih cepat yaitu 29.30 detik. Sedangkan jika prioritas utamanya adalah meminimalkan *False Negative* secara absolut untuk aplikasi klinis yang sangat sensitif, maka Information Gain dengan 75% fitur menjadi pilihan yang lebih baik, karena model ini adalah satu-satunya konfigurasi yang berhasil mencapai Recall 100%, meskipun dengan biaya waktu komputasi yang jauh lebih tinggi yaitu 80.21 detik.

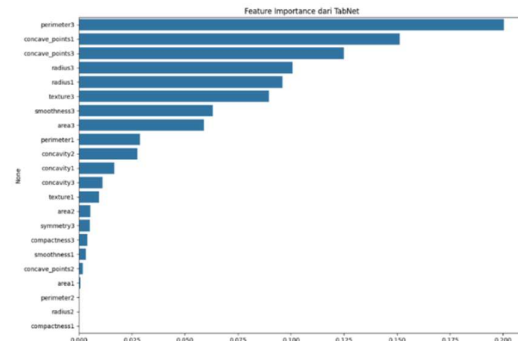
#### 4.3 Pengaruh Seleksi Fitur terhadap Feature Importance TabNet

Untuk mendapatkan wawasan yang lebih dalam tentang bagaimana seleksi fitur eksternal memengaruhi perilaku internal model, dilakukan analisis terhadap *feature importance* yang dihasilkan oleh TabNet dengan optimisasi pada tiga kondisi model terbaik. Gambar 6 menyajikan *feature importance* dari model TabNet pada kondisi dasar, yaitu ketika dilatih menggunakan seluruh 30 fitur.



Gambar 6. Feature Importance TabNet Tanpa Seleksi Fitur

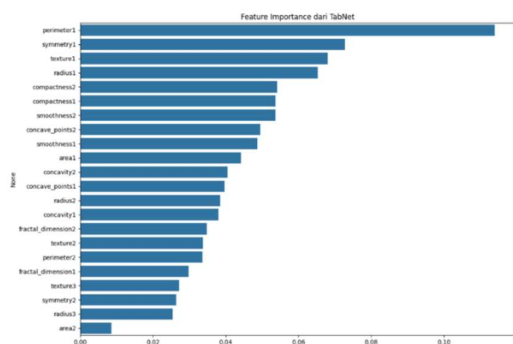
Berdasarkan Gambar 6, terlihat bahwa tanpa bantuan seleksi fitur eksternal, TabNet secara alami sudah mampu mengidentifikasi dan memprioritaskan fitur-fitur tertentu. Terdapat kecenderungan yang jelas untuk fokus pada fitur-fitur pengukuran *worst-case* (3), dengan *perimeter3* dan *concave\_points1* menempati peringkat teratas. Selanjutnya, Gambar 7 menunjukkan bagaimana distribusi kepentingan ini berubah setelah penerapan seleksi fitur Information Gain.



Gambar 7. Feature Importance TabNet dengan Seleksi Fitur Information Gain (75%)



Ketika seleksi fitur Information Gain diterapkan, kecenderungan alami TabNet ini diperkuat. Seperti yang terlihat pada Gambar 7, model menjadi lebih fokus pada subset fitur yang sama dengan model dasar. Kepentingan fitur *perimeter3* menjadi sangat dominan dengan skor sekitar 0.20, jauh lebih tinggi daripada fitur lainnya. Hal ini mengindikasikan bahwa Information Gain dan mekanisme perhatian internal TabNet memiliki pandangan yang serupa tentang fitur mana yang paling penting. Terakhir, Gambar 8 menyajikan distribusi kepentingan fitur untuk skenario Chi-Square, yang menunjukkan perubahan perilaku yang signifikan.



Gambar 8. *Feature Importance* TabNet dengan Seleksi Fitur Chi-Square (75%)

Berbeda dengan dua kondisi sebelumnya, penerapan seleksi fitur Chi-Square secara fundamental mengubah fokus perhatian TabNet. Gambar 8 menunjukkan bahwa model ini sekarang memberikan bobot kepentingan yang jauh lebih tinggi pada fitur-fitur pengukuran *mean(\_1)* dan *standard error(\_2)*, seperti *perimeter1* dan *symmetry1*. Selain itu, distribusi kepentingan fitur menjadi jauh lebih merata, tanpa ada satu pun fitur yang sangat menonjol. Perubahan fokus dan distribusi yang lebih seimbang ini dapat menjadi penjelasan mengapa model yang dilatih dengan fitur hasil seleksi Chi-Square menunjukkan perilaku yang lebih stabil dan presisi yang lebih tinggi dalam evaluasi kinerja.

#### 4.4 Perbandingan dengan Penelitian Sebelumnya

Untuk mengkontekstualisasikan temuan dari penelitian ini, kinerja model yang dicapai dibandingkan dengan hasil dari beberapa penelitian sebelumnya yang juga menggunakan dataset WDBC. Tabel 7 menyajikan perbandingan akurasi antara model TabNet yang dioptimalkan dalam penelitian ini dengan beberapa model dari penelitian terdahulu.

Tabel 7. Perbandingan Akurasi dengan Penelitian Terdahulu

Referensi	Model Utama	Akurasi
Zhou et al.	AdaBoost-Logistic	99.12%
Albadr et al.	Fast Learning Network (FLN)	98.83%
Penelitian Ini	TabNet (dioptimalkan + seleksi fitur)	98.25%
Chen et al.	XGBoost	97.40%
Ara et al.	Random Forest, SVM	96.50%

Berdasarkan perbandingan pada Tabel 7, terlihat bahwa akurasi puncak sebesar 98.25% yang dicapai dalam penelitian ini sangat

kompetitif. Hasil ini mampu mengungguli beberapa model populer seperti XGBoost (97.4%) dari Chen et al. dan Random Forest (96.5%) dari Ara et al., serta menunjukkan performa yang sangat mendekati hasil dari Albadr et al. (98.83%) dan Zhou et al. (99.12%) [7]-[10].

Namun, keunggulan utama dari penelitian ini tidak hanya terletak pada pencapaian akurasi, melainkan pada kedalaman analisis metodologisnya. Kontribusi utama dari penelitian ini adalah analisis komparatif efektivitas seleksi fitur dan analisis interpretabilitas internal model menggunakan TabNet. Pendekatan ini menawarkan wawasan yang lebih dalam dibandingkan dengan penelitian perbandingan yang umumnya hanya berfokus pada performa prediktif. Sementara beberapa studi lain melakukan seleksi fitur, penelitian ini secara spesifik membandingkan dua metode dan menganalisis bagaimana pilihan tersebut memengaruhi perilaku internal model TabNet.

Selain itu, robustitas metodologi penelitian ini juga didukung oleh proses optimisasi hiperparameter yang komprehensif dan sistematis menggunakan Optuna. Pendekatan ini memastikan bahwa perbandingan antar skenario dilakukan pada kondisi model yang paling optimal, berbeda dengan beberapa studi lain yang melaporkan hasil dengan parameter *default* atau optimisasi yang lebih terbatas. Dengan demikian, keunggulan penelitian ini tidak hanya diukur dari akurasi absolut, tetapi dari kontribusinya dalam menyajikan alur kerja yang lebih robust dan wawasan interpretabilitas yang lebih mendalam, yang merupakan aspek krusial untuk aplikasi medis.

## 5. KESIMPULAN

Berdasarkan hasil dan analisis dari penelitian yang dilakukan, dapat disimpulkan bahwa alur kerja yang mengintegrasikan seleksi fitur dan optimisasi hiperparameter secara efektif meningkatkan kinerja klasifikasi model TabNet pada dataset WDBC. Penerapan seleksi fitur terbukti mampu meningkatkan akurasi model dibandingkan dengan menggunakan seluruh 30 fitur. Dalam perbandingan metode, ditemukan adanya *trade-off* penting antara Chi-Square dan Information Gain yang bergantung pada prioritas evaluasi. Chi-Square terbukti menjadi metode yang lebih unggul untuk mencapai keseimbangan antara performa prediktif dan efisiensi komputasi. Konfigurasi optimalnya dengan 75% fitur berhasil mencapai akurasi 98.25% dan presisi 100% dengan waktu komputasi yang efisien. Di sisi lain, Information Gain menunjukkan keunggulan untuk aplikasi yang memprioritaskan sensitivitas klinis. Konfigurasinya dengan 75% fitur merupakan satu-satunya yang berhasil mencapai recall 100%, yang berarti mampu mengeliminasi risiko *False Negative*, meskipun dengan biaya presisi yang sedikit lebih rendah dan waktu komputasi yang lebih lama. Temuan ini menggarisbawahi bahwa tidak ada satu metode seleksi fitur yang terbaik secara absolut, melainkan pilihan metode yang paling efektif bergantung pada tujuan spesifik dari aplikasi. Keberhasilan penerapan seleksi fitur ini juga sangat ditentukan oleh proses optimisasi hiperparameter menggunakan Optuna, yang terbukti menjadi langkah esensial untuk membuka potensi dari subset fitur yang lebih kecil.

Temuan dan keterbatasan dari penelitian ini membuka beberapa arah untuk penelitian di masa depan. Penelitian selanjutnya disarankan untuk memvalidasi alur kerja yang diusulkan pada dataset kanker payudara lain yang lebih besar dan beragam guna menguji generalisasi dari temuan yang ada. Selain itu, eksplorasi metode seleksi fitur dari kategori lain, seperti metode *wrapper* atau *embedded*, dapat dilakukan untuk membandingkan efektivitasnya dalam mengoptimalkan performa model TabNet lebih lanjut.

## DAFTAR PUSTAKA

- [1] F. Bray *et al.*, “Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *C.A. Cancer J. Clin.*, vol. 74, pp. 229–263, May 2024, doi: [10.3322/caac.21834](https://doi.org/10.3322/caac.21834).
- [2] M. Fu, Z. Peng, M. Wu, D. Lv, Y. Li, and S. Lyu, “Current and future burden of breast cancer in Asia: A GLOBOCAN data analysis for 2022 and 2050,” *Breast*, vol. 79, no. May 2024, p. 103835, 2025, doi: [10.1016/j.breast.2024.103835](https://doi.org/10.1016/j.breast.2024.103835).
- [3] L. Pratiwi, A. Ambarsari, B. Fajri, and W. Mariyana, *Mengenal Mencegah Kanker Payudara : Sudut Pandang Teori & Penelitian*. 2024.
- [4] W. Gautama, “Breast Cancer in Indonesia in 2022: 30 Years of Marching in Place,” *Indones. J. Cancer*, vol. 16, no. 1, p. 1, 2022, doi: [10.33371/ijoc.v16i1.920](https://doi.org/10.33371/ijoc.v16i1.920).
- [5] T. S. and R. P. P. Tumuluru, C. P. Lakshmi, “A Review of Machine Learning Techniques for Breast Cancer Diagnosis in Medical Applications,” *2019 Third Int. Conf. I-SMAC (IoT Soc. Mobile, Anal. Cloud)*, vol. 11, no. 113, pp. 13–21, 2019.
- [6] and W. S. W. Wolberg, O. Mangasarian, N. Street, “Breast Cancer Wisconsin (Diagnostic).” UCI Machine Learning Repository, 1993. [Online]. Available: <https://doi.org/10.24432/C5DW2B>.
- [7] H. Chen, N. Wang, X. Du, K. Mei, Y. Zhou, and G. Cai, “Classification Prediction of Breast Cancer Based on Machine Learning,” *Comput. Intell. Neurosci.*, vol. 2023, no. 1, 2023, doi: [10.1155/2023/6530719](https://doi.org/10.1155/2023/6530719).
- [8] S. Ara, A. Das, and A. Dey, “Malignant and Benign Breast Cancer Classification using Machine Learning Algorithms,” *2021 Int. Conf. Artif. Intell. ICAI 2021*, pp. 97–101, 2021, doi: [10.1109/ICAIS2203.2021.9445249](https://doi.org/10.1109/ICAIS2203.2021.9445249).
- [9] M. A. A. Albadr, M. Ayob, S. Tiun, F. T. AL-Dhief, A. Arram, and S. Khalaf, “Breast cancer diagnosis using the fast learning network algorithm,” *Front. Oncol.*, vol. 13, no. April, pp. 1–16, 2023, doi: [10.3389/fonc.2023.1150840](https://doi.org/10.3389/fonc.2023.1150840).
- [10] S. Zhou, C. Hu, S. Wei, and X. Yan, “Breast Cancer Prediction Based on Multiple Machine Learning Algorithms,” *Technol. Cancer Res. Treat.*, vol. 23, 2024, doi: [10.1177/15330338241234791](https://doi.org/10.1177/15330338241234791).
- [11] C. Shah, Q. Du, and Y. Xu, “Enhanced TabNet: Attentive Interpretable Tabular Learning for Hyperspectral Image Classification,” *Remote Sens.*, vol. 14, no. 3, pp. 1–21, 2022, doi: [10.3390/rs14030716](https://doi.org/10.3390/rs14030716).
- [12] J. Si, W. Y. Cheng, M. Cooper, and R. G. Krishnan, “InterpreTabNet: Distilling Predictive Signals from Tabular Data by Salient Feature Interpretation,” in *Proceedings of Machine Learning Research*, ML Research Press, 2024, pp. 45353–45405.
- [13] K. Qu, J. Xu, Q. Hou, K. Qu, and Y. Sun, “Feature selection using Information Gain and decision information in neighborhood decision system,” *Appl. Soft Comput.*, vol. 136, p. 110100, 2023, doi: [10.1016/j.asoc.2023.110100](https://doi.org/10.1016/j.asoc.2023.110100).
- [14] K. Kanti Ghosh *et al.*, “Theoretical and empirical analysis of filter ranking methods: Experimental study on benchmark DNA microarray data,” *Expert Syst. Appl.*, vol. 169, no. May 2020, p. 114485, 2021, doi: [10.1016/j.eswa.2020.114485](https://doi.org/10.1016/j.eswa.2020.114485).
- [15] I. Chhillar and A. Singh, “An improved soft voting-based machine learning technique to detect breast cancer utilizing effective feature selection and SMOTE-ENN class balancing,” *Discov. Artif. Intell.*, vol. 5, no. 1, 2025, doi: [10.1007/s44163-025-00224-w](https://doi.org/10.1007/s44163-025-00224-w).
- [16] M. Bahrami, M. Vali, and H. Kia, “Breast Cancer Detection from Imbalanced Clinical Data: A Comparative Study of Sampling Methods,” *2023 30th Natl. 8th Int. Iran. Conf. Biomed. Eng. ICBME 2023*, no. December, pp. 145–149, 2023, doi: [10.1109/ICBME61513.2023.10488624](https://doi.org/10.1109/ICBME61513.2023.10488624).
- [17] R. Bhuvanya, T. Kujani, S. Manoj Kumaran, and N. Lokesh Kumar, “OptNet: Innovative Model for Early Lung Cancer Diagnosis integrating TabNet and Optuna,” *IEEE Int. Conf. Electron. Syst. Intell. Comput. ICESIC 2024 - Proc.*, pp. 174–179, 2024, doi: [10.1109/ICESIC61777.2024.10846378](https://doi.org/10.1109/ICESIC61777.2024.10846378).
- [18] A. Rahmadeyan and M. Mustakim, “Seleksi Fitur pada Supervised Learning: Klasifikasi Prestasi Belajar Mahasiswa Saat dan Pasca Pandemi COVID-19,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 1, pp. 21–32, 2023, doi: [10.25077/teknosi.v9i1.2023.21-32](https://doi.org/10.25077/teknosi.v9i1.2023.21-32).
- [19] H. I. Mun and W. Son, “Properties of chi-square statistic and information gain for feature selection of imbalanced text data,” *Korean J. Appl. Stat.*, vol. 35, no. 4, pp. 469–484, 2022, doi: [10.5351/kjas.2022.35.4.469](https://doi.org/10.5351/kjas.2022.35.4.469).
- [20] V. Borisov, T. Leemann, K. Sebler, J. Haug, M. Pawelczyk, and G. Kasneci, “Deep Neural Networks and Tabular Data: A Survey,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 35, no. 6, pp. 7499–7519, 2024, doi: [10.1109/TNNLS.2022.3229161](https://doi.org/10.1109/TNNLS.2022.3229161).
- [21] S. Arık and T. Pfister, “TabNet: Attentive Interpretable Tabular Learning,” *35th AAAI Conf. Artif. Intell. AAAI 2021*, vol. 8A, pp. 6679–6687, 2021, doi: [10.1609/aaai.v35i8.16826](https://doi.org/10.1609/aaai.v35i8.16826).
- [22] M. F. Amin, “Confusion matrix in three-class classification problems: A step-by-step tutorial,” *J. Eng. Res.*, vol. 6, no. 5, 2023, [Online]. Available: [https://erjeng.journals.ekb.eg/article/296718\\_30a98aac15193d04dc73ba9bc00cf046.pdf](https://erjeng.journals.ekb.eg/article/296718_30a98aac15193d04dc73ba9bc00cf046.pdf)