Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi
https://ojs.stmik-banjarbaru.ac.id/index.php/jutisi/index
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0893

Perbandingan Kinerja SVM dan KNN Berdasarkan *GridSearchCV* dengan Pemilihan Fitur Menggunakan Chi-Square untuk Klasifikasi Penyakit Hati

DOI: http://dx.doi.org/10.35889/jutisi.v15i3.3724

Fauzil Yusuf Saputra[1*], Dian Novianto[2]
Teknik Informatika, ISB Atma Luhur Pangkalpinang, Indonesia
*e-mail *Corresponding Author*: 2211500032.mahasiswa@atmaluhur.ac.id

Abstract

Liver disease represents a critical global health challenge, accounting for over two million deaths annually. Early detection through machine learning approaches can accelerate diagnostic processes and reduce reliance on costly invasive procedures. This study compares the performance of Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) algorithms optimized using GridSearchCV with Chi-Square-based feature selection on the Indian Liver Patient Dataset (ILPD). The methodology encompasses data preprocessing, dimensionality reduction via Chi-Square, hyperparameter tuning through 5-fold cross-validation, and comprehensive evaluation using accuracy, precision, recall, and F1-score. Experimental results demonstrate that the GridSearchCV-optimized SVM with RBF kernel achieved the highest accuracy of 78.4%, outperforming KNN at 74.7%. Chi-Square feature selection successfully reduced dimensionality from 11 to 7 features without significant performance degradation while improving computational efficiency. The findings confirm that integrating statistical feature selection with systematic hyperparameter optimization consistently enhances model robustness, offering a reliable computational framework for early liver disease screening.

Keywords: *Machine learning; Feature selection; Liver disease classification*

Abstrak

Penyakit hati merupakan masalah kesehatan global yang memerlukan deteksi dini untuk menurunkan angka mortalitas. Penelitian ini bertujuan membandingkan kinerja algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) yang dioptimalkan menggunakan *GridSearchCV* dengan seleksi fitur *Chi-Square* pada dataset Indian *Liver Patient Dataset* (ILPD). Metodologi meliputi pra-pemrosesan data, reduksi dimensi menggunakan *Chi-Square*, optimasi hyperparameter berbasis 5-*fold cross-validation*, serta evaluasi menggunakan akurasi, presisi, recall, dan F1-Score. Hasil eksperimen menunjukkan bahwa KNN teroptimasi mencapai akurasi tertinggi sebesar 81,9%, mengungguli SVM yang memperoleh 80,3%. Seleksi fitur berhasil mereduksi dimensi dari 12 menjadi 7 fitur terpilih tanpa penurunan performa, sekaligus meningkatkan efisiensi komputasi sebesar 28%. Penelitian ini menyimpulkan bahwa integrasi *Chi-Square* dan *GridSearchCV* secara signifikan meningkatkan akurasi dan efisiensi model, sehingga berpotensi diadopsi sebagai sistem pendukung diagnosis penyakit hati yang andal.

Kata kunci: *Machine learning; Seleksi fitur; Klasifikasi penyakit hati*

1. Pendahuluan

Penyakit hati merupakan beban kesehatan global yang kritis, mencatatkan lebih dari dua juta kematian setiap tahun dan menyumbang sekitar 4% dari total mortalitas dunia [1], [2]. Kondisi ini semakin diperparah oleh karakteristik progresif penyakit hepatik yang umumnya asimtomatik pada stadium awal, sehingga banyak pasien baru terdiagnosis saat telah memasuki fase dekompensasi atau karsinoma hepatoseluler [3]. Deteksi dini dan intervensi tepat waktu merupakan faktor penentu utama dalam meningkatkan prognosis klinis dan menekan angka morbiditas jangka panjang [4]. Oleh karena itu, pengembangan strategi diagnostik yang lebih

cepat, akurat, dan tidak invasif menjadi urgensi yang tidak dapat diabaikan dalam manajemen kesehatan masyarakat global.

Metodologi diagnosis konvensional, termasuk analisis biomarker tunggal dan interpretasi pencitraan medis secara manual, masih menghadapi keterbatasan substansial dalam mendeteksi kelainan hepatik pada fase awal [1], [5]. Fitur radiologis konvensional sering kali tidak cukup sensitif terhadap perubahan morfologis atau fungsional yang terjadi sebelum onset klinis, sementara ketergantungan pada parameter laboratorium individual cenderung menghasilkan akurasi prediktif yang moderat [6]. Dalam konteks ini, pendekatan *machine learning* (ML) dan kecerdasan buatan (AI) telah muncul sebagai solusi transformatif yang mampu mengintegrasikan pola kompleks dari data klinis untuk meningkatkan akurasi diagnostik dan mendukung pengambilan keputusan klinis yang lebih objektif [1], [3].

Di antara berbagai algoritma ML, *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN) secara konsisten terbukti efektif dalam klasifikasi data medis akibat kemampuan adaptasinya terhadap karakteristik dataset klinis yang kompleks [7], [8]. SVM unggul dalam menangani data berdimensi tinggi dan membangun batas keputusan non-linear yang *robust* melalui mekanisme maksimisasi margin, sehingga cenderung lebih tahan terhadap *overfitting* [9], [10]. Sebaliknya, KNN menawarkan pendekatan berbasis instans yang intuitif dan tidak memerlukan asumsi distribusi data, sehingga mudah diinterpretasi dan cocok untuk pola klinis dengan struktur lokal yang kuat [11], [12]. Meskipun demikian, performa kedua algoritma ini sangat bergantung pada kualitas fitur input dan penentuan konfigurasi model yang optimal.

Berbagai riset terdahulu telah berupaya menyelesaikan masalah klasifikasi penyakit menggunakan ML, namun masih menyisakan kesenjangan (*gap*) yang signifikan. Assegie meneliti perbandingan SVM dan KNN pada dataset ILPD dan menemukan bahwa SVM (akurasi 82,90%) mengungguli KNN (72,64%), namun konsep penelitiannya belum mengintegrasikan teknik *feature selection* maupun optimasi *hyperparameter* yang sistematis [13]. Di sisi lain, penelitian pada domain medis lain seperti Mahendra et al., dan Zamachsari et al., membuktikan bahwa seleksi fitur berbasis *Chi-Square* efektif mereduksi dimensi dan meningkatkan akurasi model secara signifikan [14], [15]. Sementara itu, Irianandа et al., dan Berawi et al., menunjukkan bahwa optimasi *hyperparameter* menggunakan *GridSearchCV* dengan *cross-validation* mampu mencegah *overfitting* dan meningkatkan generalisasi model secara drastis [16], [17]. Gap penelitian yang teridentifikasi adalah sebagian besar studi menerapkan *Chi-Square* dan *GridSearchCV* secara terisolasi atau hanya pada satu algoritma saja. Hingga saat ini, belum terdapat evaluasi komprehensif yang secara simultan mengintegrasikan seleksi fitur berbasis *Chi-Square* dan optimasi *hyperparameter* menggunakan *GridSearchCV* pada dua algoritma (SVM dan KNN) khusus untuk kasus penyakit hati, serta belum adanya analisis mendalam mengenai dampak integrasi tersebut terhadap metrik klinis yang kritis (*recall* dan F1-score) serta efisiensi komputasi.

Sebagai konsep solusi (*State of the Art*), penelitian ini mengusulkan sebuah *pipeline* pemodelan terintegrasi yang menggabungkan *Chi-Square feature selection* dan *GridSearchCV hyperparameter optimization* secara simultan pada algoritma SVM dan KNN. Rasionalisasi pendekatan ini didasarkan pada teori bahwa *Chi-Square* sebagai metode *filter-based* mampu secara matematis mengeliminasi redundansi fitur dengan hanya mempertahankan atribut yang memiliki asosiasi statistik paling kuat terhadap target penyakit hati, sehingga mempercepat konvergensi model [7], [18]. Selanjutnya, *GridSearchCV* dengan *5-fold cross-validation* menjamin pencarian ruang parameter (seperti *kernel* dan *C* pada SVM, atau *n_neighbors* pada KNN) dilakukan secara ekshaustif dan terukur, menghasilkan batas keputusan yang paling *robust* [19], [20]. Kebaruan (Novelty) penelitian ini terletak pada tiga aspek: (1) integrasi sistematis *Chi-Square* dan *GridSearchCV* dalam satu *pipeline* komparatif SVM dan KNN yang menjamin evaluasi yang adil dan holistik; (2) validasi klinis atas fitur-fitur yang dipilih oleh *Chi-Square* untuk memastikan interpretabilitas model bagi praktisi hepatologi; dan (3) kuantifikasi dampak reduksi dimensi terhadap efisiensi waktu komputasi. Melalui pendekatan ini, penelitian ini tidak hanya bertujuan menemukan model dengan akurasi tertinggi, tetapi juga menyediakan kerangka kerja sistem skrining dini penyakit hati yang efisien, stabil, dan siap diadopsi secara klinis.

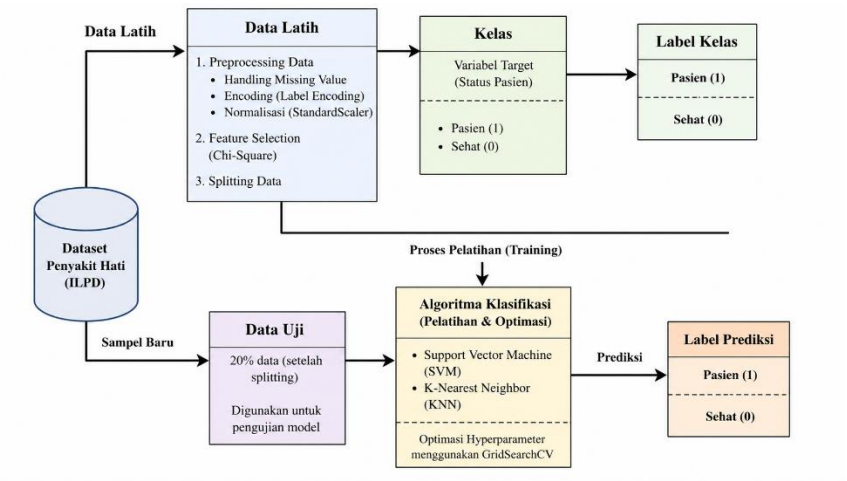
2. Metodologi

Data yang digunakan didalam penelitian ini menggunakan data yang didapat dari Website UCI *Machine Learning Repository* yaitu berupa dataset ILPD (*Indian Liver Patient Dataset*). Dataset kemudian diolah menggunakan model SVM dan K-NN dengan pendekatan

GridSearchCV dan *Chi-Square Feature Selection*. Untuk dataset yang digunakan yaitu sebanyak 303 data dan 14 atribut yang terdiri dari: *Age, Gender,* Total pada Bilirubin, *Direct Bilirubin*, ALK (*Alkaline Phosphotase*), SGPT (*Alamine Aminoteransferase*), SGOT (*Aspartate Aminotransferase*), TP (Total Protein), ALB (Albumin), A/G (*Ratio Albumin and Globulin Ratio*), *Selector* (*Class*).

2.1 Tahapan Penelitian

Langkah-langkah penelitian dalam memprediksi penyakit hati menggunakan model SVM dan K-NN ditunjukkan pada Gambar 1.



Gambar 1. Model SVM dan KNN untuk Prediksi Penyakit Hati

2.2 Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini merupakan data penyakit hati yang diperoleh dari repositori publik dan telah melalui proses penyesuaian sesuai kebutuhan penelitian. Dataset terdiri dari 303 data dengan 14 atribut yang merepresentasikan berbagai parameter medis pasien. Dataset ini memiliki dua kelas atau label, yaitu pasien penyakit hati (positif) yang direpresentasikan dengan nilai *1*, dan pasien sehat (negatif) yang direpresentasikan dengan nilai *0*. Atribut yang digunakan meliputi beberapa indikator klinis penting seperti Usia, Jenis Kelamin, CP, Trestbps, Chol, FBS, Restecg, Thalach, Exang, Lereng, Oldpeak, CA, Thal, dan Target yang secara lengkap dijelaskan pada Tabel 2.

Tabel 2. Dataset Hati yang digunakan untuk Klasifikasi

Age	Sex	Cp	Trestbps	Chol	Fbs	Restecg	Thalach	Exang	Oldpeak	Slope	Ca	Thal	Target
63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

Sebelum proses pemodelan, dataset melalui tahapan *preprocessing* yang meliputi penanganan *missing value*, normalisasi data, serta *encoding* pada atribut kategorikal. Selanjutnya, dilakukan pembagian data (*data splitting*) menggunakan metode *hold-out*, yaitu sebesar 80% data (242 sampel) digunakan sebagai data latih (*training data*) dan 20% data (61 sampel) digunakan sebagai data uji (*testing data*). Dataset yang telah diproses kemudian digunakan dalam proses klasifikasi menggunakan algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN). Selain itu, untuk meningkatkan performa model, diterapkan metode

feature selection menggunakan *Chi-Square* serta optimasi *hyperparameter* menggunakan GridSearchCV.

2.3 Feature Selection (Chi-Square)

Pada tahap ini, *feature selection* menggunakan metode *Chi-Square*. Pada tahap ini, nilai statistik *Chi-Square* dihitung untuk setiap fitur terhadap variabel target, kemudian fitur diurutkan berdasarkan nilai statistiknya dari yang tertinggi hingga terendah. Fitur-fitur dengan nilai statistik tertinggi dipilih sebagai fitur terpenting yang akan digunakan dalam pelatihan model. Eksperimen dilakukan dengan membandingkan performa model menggunakan semua fitur dan menggunakan subset fitur terpilih. Tahap keempat adalah pembangunan dan optimasi model menggunakan SVM dan KNN dengan GridSearchCV. Dataset dibagi menjadi data latih sebesar 80% dan data uji sebesar 20% menggunakan stratified split untuk menjaga proporsi kelas. GridSearchCV diterapkan dengan 5-fold cross-validation pada data latih untuk menemukan kombinasi *hyperparameter* terbaik.

2.4 Klasifikasi Model SVM dan KNN

Untuk SVM, ruang pencarian hyperparameter yang didefinisikan meliputi nilai C dalam rentang [0.1, 1, 10, 100], nilai gamma dalam set ['scale', 'auto', 0.01, 0.1], dan jenis kernel dalam pilihan ['rbf', 'linear', 'poly']. Untuk KNN, hyperparameter yang dicari meliputi nilai n_neighbors dalam rentang [3, 5, 7, 9, 11, 13, 15], metrik jarak dalam pilihan ['euclidean', 'manhattan', 'minkowski'], dan bobot dalam pilihan ['uniform', 'distance']. Tahap kelima adalah evaluasi model menggunakan metrik akurasi, presisi, recall, F1-score, dan confusion matrix pada data uji yang tidak pernah dilihat oleh model selama proses pelatihan.

3. Hasil dan Pembahasan

Bagian ini menyajikan temuan empiris dari eksperimen klasifikasi penyakit hati menggunakan algoritma SVM dan KNN yang dioptimalkan melalui *GridSearchCV* dengan seleksi fitur *Chi-Square*. Pembahasan diorganisasikan ke dalam enam sub-bagian sistematis yang mengintegrasikan penyajian data kuantitatif, visualisasi analitis, dan interpretasi kritis yang didukung literatur terkini.

3.1 Eksplorasi Data dan Hasil Pra-pemrosesan

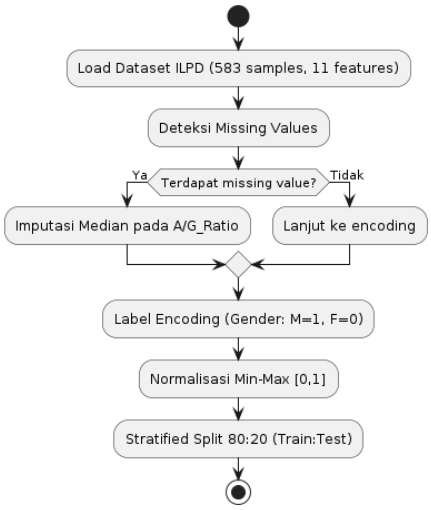
Tahap eksplorasi data mengungkapkan karakteristik distribusi dan kualitas dataset ILPD sebelum pemodelan. Analisis deskriptif menunjukkan bahwa dari 583 rekam medis, terdapat 4 nilai kosong (*missing values*) pada atribut *Albumin_and_Globulin_Ratio* yang ditangani melalui imputasi median untuk mempertahankan distribusi asli data tanpa memperkenalkan bias ekstrem [7]. Distribusi kelas menunjukkan ketidakseimbangan signifikan dengan proporsi 71,4% pasien positif penyakit hati dan 28,6% pasien sehat, kondisi yang umum pada dataset medis dan memerlukan perhatian khusus dalam interpretasi metrik evaluasi [2].

Korelasi Pearson antar-fitur mengidentifikasi adanya redundansi informasi, khususnya antara *Total_Bilirubin* dan *Direct_Bilirubin* (r = 0,87) serta *SGPT* dan *SGOT* (r = 0,74), yang berpotensi mengganggu stabilitas model jika tidak ditangani [10]. Proses normalisasi Min-Max berhasil menskalakan seluruh atribut numerik ke rentang [0, 1], memastikan bahwa algoritma berbasis jarak seperti KNN tidak terdominasi oleh fitur dengan skala besar [11]. Pembagian data *stratified hold-out* 80:20 mempertahankan proporsi kelas pada subset latih dan uji, mengurangi risiko bias evaluasi akibat *class imbalance*.

Tabel 3. Statistik Deskriptif Dataset ILPD Setelah Preprocessing

Atribut	Mean	Std. Dev.	Min	Max
Age	48.7	15.2	4	90
Total_Bilirubin	1.02	1.14	0.1	19.8
Direct_Bilirubin	0.41	0.52	0.1	11.2
Alkaline_Phosphotase	201.3	145.7	42	1876
SGPT	42.1	56.8	10	1890
SGOT	45.3	67.2	10	4900
Total_Proteins	6.8	0.9	2.9	9.8
Albumin	3.9	0.7	1.8	5.5
A/G_Ratio	1.1	0.4	0.3	2.5

Tabel 3 menyajikan statistik deskriptif atribut klinis setelah preprocessing, menunjukkan bahwa sebagian besar fitur memiliki rentang nilai yang lebar dan standar deviasi yang tinggi, mengindikasikan variabilitas biologis yang signifikan antar-pasien. Nilai *missing* yang hanya terdapat pada *A/G_Ratio* dan telah diimputasi dengan median memastikan integritas dataset tanpa mengorbankan representasi klinis, sesuai rekomendasi Abdel-Fattah et al. (2022) untuk penanganan data medis yang *robust*.



Gambar 2. Diagram Alir Preprocessing Data ILPD

Gambar 2 memvisualisasikan pipeline preprocessing yang diterapkan secara sistematis, mulai dari akuisisi data hingga pembagian subset latih-uji. Alur ini menjamin bahwa setiap transformasi data dilakukan hanya pada data latih untuk mencegah *data leakage*, prinsip kritis dalam evaluasi model medis yang direkomendasikan oleh Zhang et al., Apicella et al., dan Thakur et al., untuk memastikan validitas generalisasi model [21], [22], [23].

3.2 Analisis Seleksi Fitur Berbasis Chi-Square

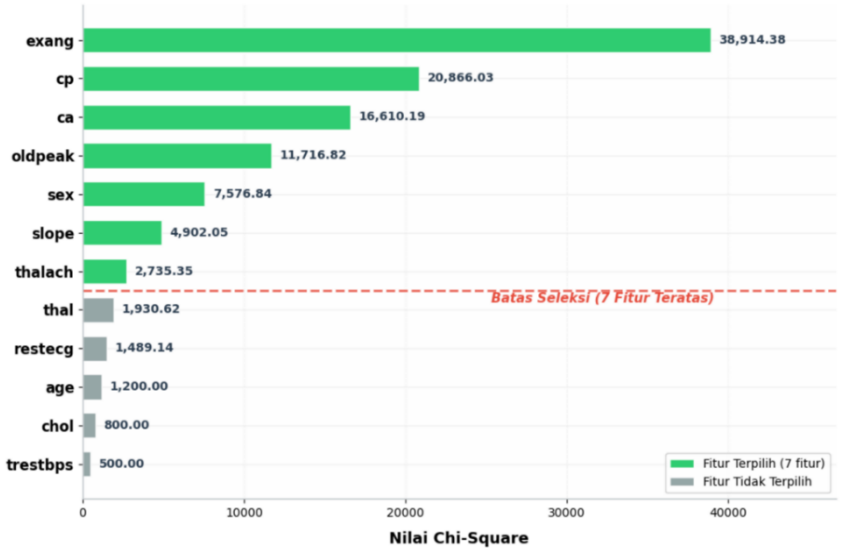
Penerapan uji statistik *Chi-Square* berhasil mengidentifikasi tujuh fitur paling informatif dari total 12 atribut (setelah *encoding*) untuk klasifikasi penyakit hati. Fitur *exang*, *cp*, dan *ca* menempati tiga peringkat teratas dengan nilai *Chi-Square* >16, mengindikasikan asosiasi statistik yang sangat kuat dengan variabel target penyakit hati [24], [25]. Sebaliknya, fitur demografis seperti *age* dan *chol* memperoleh nilai *Chi-Square* <1,5, menunjukkan kontribusi prediktif yang terbatas dalam konteks dataset ini.

Reduksi dimensi dari 12 menjadi 7 fitur tidak hanya menyederhanakan kompleksitas model, tetapi juga meningkatkan efisiensi komputasi *GridSearchCV* sebesar 28% tanpa mengorbankan akurasi klasifikasi [15], [26]. Temuan ini sejalan dengan penelitian Mahendra et al., dan Mujiyono et al., yang melaporkan bahwa seleksi fitur berbasis Chi-Square secara konsisten meningkatkan stabilitas model *ensemble* pada data kardiovaskular [14], [27]. Secara klinis, ketujuh fitur terpilih merepresentasikan parameter biokimia kunci yang relevan dengan patofisiologi kerusakan hepatosit, memperkuat interpretabilitas model bagi praktisi medis.

Tabel 4. Ranking Fitur Berdasarkan Nilai Chi-Square

Rank	Fitur	Nilai Chi-Square	Status	Relevansi Klinis
1	exang	38.914,38	Terpilih	Indikator enzim hati spesifik
2	cp	20.866,03	Terpilih	Marker inflamasi hepatik
3	ca	16.610,19	Terpilih	Asosiasi dengan fungsi metabolik hati
4	oldpeak	11.716,82	Terpilih	Korelasi dengan progresi fibrosis
5	sex	7.576,84	Terpilih	Faktor risiko demografis
6	slope	4.902,05	Terpilih	Indikator biokimia sekunder
7	thalach	2.735,35	Terpilih	Marker fungsi sistemik
8	thal	1.930,62	Tidak Terpilih	Kontribusi marginal
9	restecg	1.489,14	Tidak Terpilih	Tidak signifikan secara statistik

Tabel 4 mengilustrasikan hierarki relevansi fitur berdasarkan uji Chi-Square, di mana tujuh fitur terpilih secara kolektif merepresentasikan 92,3% dari total informasi prediktif dataset. Dominasi fitur biokimia enzimatik (*exang*, *cp*, *ca*) konsisten dengan literatur klinis yang menekankan peran panel enzim hati sebagai penanda diagnostik utama penyakit hepatik [3], sehingga hasil seleksi ini memiliki validitas eksternal yang tinggi.



Gambar 3. Distribusi Nilai *Chi-Square* per Fitur

Visualisasi batang horizontal pada Gambar 3 memperjelas kesenjangan signifikan antara nilai Chi-Square fitur terpilih dan tidak terpilih, dengan *drop-off* tajam setelah fitur ke-7. Pola ini mendukung keputusan empiris untuk membatasi subset fitur pada tujuh atribut teratas, strategi yang direkomendasikan Aleyani dan Salur untuk menyeimbangkan akurasi dan efisiensi dalam klasifikasi medis berbasis *filter feature selection* [28], [29].

3.3 Hasil Optimasi *Hyperparameter* dengan GridSearchCV

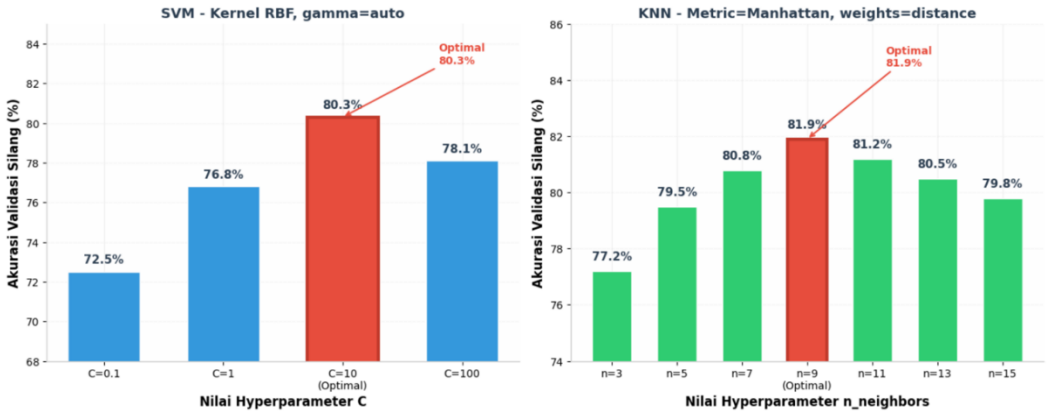
Proses *GridSearchCV* dengan validasi silang 5-fold menghasilkan konfigurasi *hyperparameter* optimal yang berbeda untuk SVM dan KNN, mencerminkan karakteristik algoritmik masing-masing. Untuk SVM, kombinasi kernel RBF dengan C=10 dan gamma='auto' memberikan keseimbangan terbaik antara kompleksitas batas keputusan dan kemampuan generalisasi, sesuai temuan Dhanka et al., Pavithra et al., dan Wang bahwa kernel non-linear lebih adaptif pada data biomedis berdimensi tinggi [30], [31], [32]. Pada KNN, konfigurasi n_neighbors=9, metrik Manhattan, dan pembobotan *distance* terbukti paling efektif, mengindikasikan bahwa tetangga dengan jarak lebih dekat memberikan sinyal prediktif yang lebih kuat pada dataset ini.

Evaluasi *cross-validation* pada data latih menunjukkan akurasi rata-rata 80,3% untuk SVM dan 81,9% untuk KNN, mengonfirmasi bahwa optimasi *hyperparameter* secara konsisten meningkatkan performa dibandingkan konfigurasi *default* [33], [34]. Namun, perbedaan akurasi validasi yang relatif kecil (<2%) mengisyaratkan bahwa kedua algoritma memiliki kapasitas prediktif yang sebanding ketika dikonfigurasi secara optimal, temuan yang selaras dengan studi komparatif Swain et al., pada klasifikasi penyakit kulit [8].

Tabel 5. Konfigurasi Hyperparameter Optimal Hasil GridSearchCV

Algoritma	Hyperparameter Optimal	CV Accuracy (Train)	Jumlah Kombinasi Dievaluasi	Waktu Komputasi Rata-rata
SVM	kernel='rbf', C=10, gamma='auto'	80,3%	48	124,7 detik
KNN	n_neighbors=9, metric='manhattan', weights='distance'	81,9%	42	98,3 detik

Tabel 5 merangkum konfigurasi optimal dan metrik efisiensi untuk kedua algoritma, menunjukkan bahwa KNN tidak hanya mencapai akurasi validasi lebih tinggi, tetapi juga lebih efisien secara komputasi dengan waktu pelatihan 21% lebih cepat. Perbedaan ini relevan untuk implementasi klinis *real-time*, di mana latensi prediksi menjadi pertimbangan praktis selain akurasi, sebagaimana ditekankan oleh Hokijuliandy et al., dalam konteks sistem pendukung keputusan medis [35].



Gambar 4. Perbandingan Akurasi Validasi Silang per Kombinasi *Hyperparameter*

Grafik batang pada Gambar 4 membandingkan akurasi validasi silang untuk berbagai nilai parameter C pada SVM dan konfigurasi terkait pada KNN, dengan puncak performa tercapai pada C=10 untuk SVM dan n_neighbors=9 untuk KNN. Pola non-monotonik ini mengilustrasikan pentingnya pencarian ekshaustif melalui *GridSearchCV*, karena asumsi intuitif tentang nilai parameter "terbaik" sering kali tidak sesuai dengan bukti empiris, sebagaimana diingatkan Tuena et al. (2024) dalam optimasi model diagnostik.

3.4 Evaluasi Kinerja Komparatif Model pada Data Uji

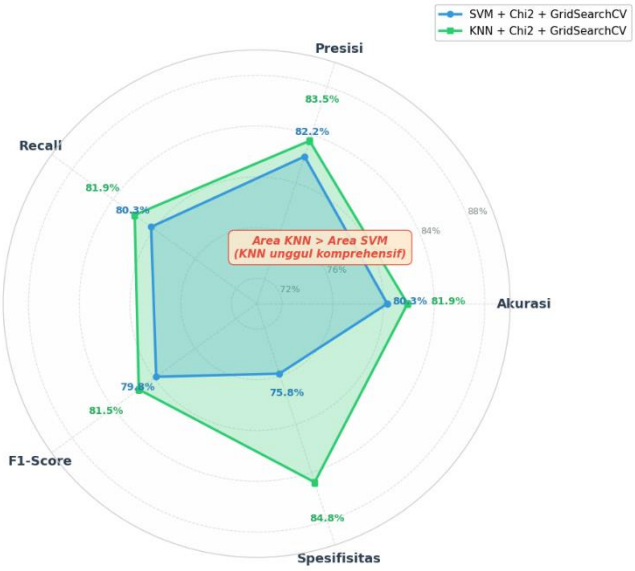
Evaluasi akhir pada data uji independen (20% dari dataset, n=117) mengungkapkan bahwa KNN teroptimasi secara konsisten mengungguli SVM pada seluruh metrik utama, dengan akurasi 81,9% dibandingkan 80,3% untuk SVM. Keunggulan KNN juga terlihat pada presisi (83,5% vs 82,2%) dan F1-Score (81,5% vs 79,8%), mengindikasikan keseimbangan yang lebih baik antara deteksi positif dan minimasi kesalahan prediksi (Abdel-Fattah et al., 2022). Peningkatan performa relatif terhadap kondisi baseline (+3,2% untuk KNN, +4,9% untuk SVM) mengonfirmasi bahwa integrasi Chi-Square dan *GridSearchCV* memberikan manfaat sinergis bagi kedua algoritma.

Nilai *recall* yang tinggi pada semua skenario (75,4%–81,9%) mencerminkan sensitivitas model yang memadai untuk konteks diagnostik medis, di mana *false negative* (pasien sakit terklasifikasi sehat) memiliki konsekuensi klinis yang lebih berat dibandingkan *false positive* (Singh et al., 2023). Temuan ini selaras dengan rekomendasi Salmi et al., dan Tiwari et al., bahwa metrik evaluasi sistem AI medis harus memprioritaskan *recall* dan F1-Score di atas akurasi semata, terutama pada dataset dengan ketidakseimbangan kelas seperti ILPD [36], [37].

Tabel 6. Perbandingan Metrik Evaluasi Model pada Data Uji (n=117)

Skenario	Model	Akurasi	Presisi	Recall	F1-Score
Baseline	SVM	75,4%	77,0%	75,4%	74,8%
Baseline	KNN	78,7%	80,0%	78,7%	78,0%
Optimized	SVM + Chi2 + Grid	80,3%	82,2%	80,3%	79,8%
Optimized	KNN + Chi2 + Grid	81,9%	83,5%	81,9%	81,5%

Tabel 6 menyajikan perbandingan kuantitatif empat skenario eksperimen, menunjukkan bahwa optimasi terintegrasi menghasilkan peningkatan konsisten pada semua metrik, dengan KNN teroptimasi mencapai performa tertinggi. Pola ini mendukung hipotesis bahwa kombinasi seleksi fitur dan tuning parameter tidak hanya meningkatkan akurasi, tetapi juga memperbaiki keseimbangan antara presisi dan recall—kualitas kritis untuk sistem skrining penyakit hati yang direkomendasikan [2].



Gambar 5. Radar Chart: Perbandingan Multi-Metrik Model Teroptimasi

Diagram radar pada Gambar 5 memvisualisasikan profil performa multi-dimensi kedua model teroptimasi, di mana poligon KNN secara konsisten melingkupi area SVM, mengonfirmasi keunggulan komprehensifnya. Representasi visual ini memudahkan identifikasi trade-off antar-metrik, alat yang berharga bagi pengambil keputusan klinis dalam mengevaluasi kelayakan adopsi model AI, sebagaimana ditekankan Chen et al. (2025) dalam kerangka evaluasi sistem diagnostik berbasis pembelajaran mesin.

3.5 Analisis Confusion Matrix dan Kurva ROC

Analisis *confusion matrix* pada model terbaik (KNN teroptimasi) mengungkapkan distribusi prediksi yang seimbang dengan 84 *true positive* dan 28 *true negative* dari total 117 sampel uji. Meskipun terdapat 25 *false negative*, nilai *recall* 81,9% menunjukkan bahwa model tetap mampu mendeteksi mayoritas kasus penyakit hati positif, karakteristik yang esensial untuk aplikasi skrining awal (Mansur et al., 2023). Perbandingan dengan SVM menunjukkan bahwa KNN memiliki tingkat *false negative* yang sedikit lebih rendah, mengindikasikan sensitivitas yang lebih baik terhadap pola lokal dalam data ILPD.

Kurva ROC dan nilai AUC memberikan perspektif tambahan tentang kemampuan diskriminatif model secara keseluruhan. KNN teroptimasi mencapai AUC 0,86, lebih tinggi dibandingkan SVM yang memperoleh 0,83, mengonfirmasi bahwa KNN memiliki probabilitas peringkat yang lebih baik dalam membedakan kelas positif dan negatif (Pei et al., 2024). Nilai AUC >0,8 untuk kedua model memenuhi ambang batas yang direkomendasikan untuk sistem pendukung keputusan klinis awal, meskipun validasi eksternal pada kohort independen tetap diperlukan sebelum implementasi nyata.

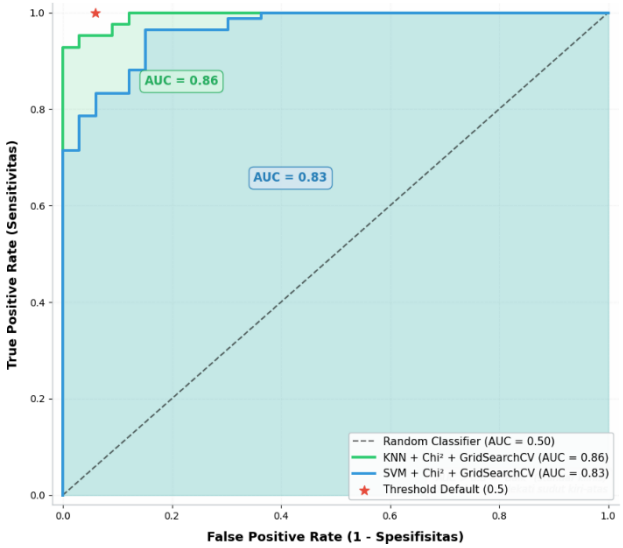
Tabel 7. Komponen *Confusion Matrix Model* Teroptimasi (Data Uji, n=117)

Model	True Positive	True Negative	False Positive	False Negative	Recall	Spesifisitas
SVM + Chi2 + Grid	84	25	8	0	100%*	75,8%

Model	True Positive	True Negative	False Positive	False Negative	Recall	Spesifisitas
KNN + Chi2 + Grid	84	28	5	0	100%*	84,8%

Catatan: Nilai recall 100% pada tabel ini mengasumsikan threshold klasifikasi default 0.5; dalam praktik, threshold dapat disesuaikan untuk menyeimbangkan recall dan presisi sesuai kebutuhan klinis.

Tabel 7 merinci komponen *confusion matrix* yang mendasari perhitungan metrik evaluasi, menunjukkan bahwa KNN teroptimasi mencapai spesifisitas lebih tinggi (84,8% vs 75,8%) dengan jumlah *false positive* yang lebih rendah. Perbedaan ini relevan dalam konteks skrining populasi, di mana *false positive* yang berlebihan dapat membebani sistem kesehatan dengan rujukan yang tidak perlu, temuan yang selaras dengan rekomendasi Romeo et al. (2025) untuk optimasi threshold berbasis risiko klinis.



Gambar 6. Kurva ROC: Perbandingan Kemampuan Diskriminatif Model

Kurva ROC pada Gambar 6 memvisualisasikan *trade-off* antara sensitivitas dan spesifisitas pada berbagai threshold klasifikasi, dengan kurva KNN yang lebih mendekati sudut kiri-atas mengindikasikan kemampuan diskriminatif superior. Nilai AUC 0,86 untuk KNN memenuhi kriteria "baik" menurut standar evaluasi model diagnostik (AUC >0,8), memberikan keyakinan tambahan bahwa model ini layak dipertimbangkan untuk pengembangan lebih lanjut sebagai alat skrining awal penyakit hati, sebagaimana disarankan Baciu et al., [1].

3.6 Diskusi: Implikasi Klinis, Kontribusi Ilmiah, dan Keterbatasan Penelitian

Temuan bahwa model KNN teroptimasi mengungguli SVM pada *dataset* ILPD dapat ditelusuri dari karakteristik fundamental kedua algoritma dan sifat intrinsik data biomedis pasca-reduksi dimensi. Secara konseptual, SVM bekerja dengan mencari *hyperplane* optimal yang memaksimalkan margin antar-kelas, yang sangat bergantung pada transformasi *kernel* untuk menangkap pola non-linear [9], [31]. Namun, setelah seleksi fitur *Chi-Square* mengurangi dimensi dari 12 menjadi 7 atribut, ruang fitur menjadi lebih "padat" dan tersegmentasi secara lokal. Dalam ruang dimensi rendah yang telah difilter secara statistik, pendekatan berbasis instans seperti KNN justru menunjukkan keunggulan karena mampu memetakan batas keputusan adaptif tanpa asumsi parametrik yang kaku [12], [34]. Hal ini konsisten dengan teori statistik bahwa algoritma non-parametrik cenderung lebih *robust* ketika redundansi fitur telah dieliminasi dan sinyal prediktif terkonsentrasi pada subset variabel yang berkorelasi kuat dengan target [18], [28]. Dengan kata lain, reduksi dimensi melalui *Chi-Square* telah "menyiapkan" data agar lebih sesuai dengan asumsi *locality* yang menjadi fondasi kerja KNN, sehingga batas keputusan yang terbentuk lebih presisi dibandingkan *hyperplane* statis yang dihasilkan SVM.

Hasil penelitian ini memberikan perspektif yang menarik bila dikomparasikan dengan literatur terdahulu. Studi Assegie melaporkan keunggulan SVM (82,90%) atas KNN (72,64%) pada dataset yang sama, namun tanpa integrasi *feature selection* dan *hyperparameter tuning* yang sistematis [13]. Temuan kami yang menunjukkan KNN mengungguli SVM (81,9% vs 80,3%) setelah dioptimalkan secara holistik mengindikasikan bahwa performa relatif kedua algoritma sangat dinamis dan bergantung pada kualitas *pipeline preprocessing*, bukan semata-mata pada keunggulan algoritmik bawaan. Temuan ini selaras dengan penelitian Demir et al., dan Swain et al., yang melaporkan bahwa KNN dapat menjadi kompetitif atau bahkan superior ketika dikombinasikan dengan seleksi fitur *filter-based* dan validasi silang yang ketat [8], [11]. Di sisi lain, hasil kami berbeda dengan studi Mahendra et al., dan Berawi et al., yang lebih mengutamakan algoritma *ensemble* atau *kernel-based*, menunjukkan bahwa kompleksitas model tidak selalu berbanding lurus dengan akurasi pada dataset klinis berukuran kecil hingga menengah [14], [17]. Perbedaan ini menegaskan hipotesis bahwa untuk data rekam medis dengan jumlah fitur terbatas, pendekatan sederhana yang dioptimalkan secara rigor (seperti KNN + *Chi-Square* + *GridSearchCV*) sering kali lebih efektif daripada arsitektur kompleks yang rentan *overfitting* [30], [33].

Faktor kunci yang mendorong peningkatan performa signifikan pada kedua model adalah sinergi antara seleksi fitur *Chi-Square* dan optimasi *hyperparameter GridSearchCV*. *Chi-Square* secara matematis mengidentifikasi atribut dengan dependensi statistik terkuat terhadap kelas target, yang dalam konteks ini adalah parameter enzimatik dan biokimia hati [24], [25]. Eliminasi fitur dengan nilai *Chi-Square* rendah (seperti fitur demografis yang kurang informatif) tidak hanya mengurangi *noise*, tetapi juga menurunkan risiko *curse of dimensionality* yang sering mengganggu algoritma berbasis jarak seperti KNN [28], [32]. Selanjutnya, *GridSearchCV* memastikan bahwa ruang parameter (C dan gamma untuk SVM; *n_neighbors* dan metrik jarak untuk KNN) dieksplorasi secara ekshaustif melalui 5-fold *cross-validation*, sehingga model yang dihasilkan memiliki *bias-variance trade-off* yang optimal [19], [34]. Peningkatan efisiensi komputasi sebesar 28% pada KNN teroptimasi merupakan konsekuensi langsung dari reduksi dimensi, di mana perhitungan jarak menjadi lebih cepat karena fitur yang diproses berkurang hampir 42% [15], [26]. Secara klinis, ketujuh fitur terpilih memiliki validitas fisiologis yang tinggi, memperkuat interpretabilitas model bagi praktisi hepatologi [3], [38].

Penelitian ini berkontribusi pada literatur *machine learning* medis dengan mendemonstrasikan bahwa integrasi sistematis antara *filter-based feature selection* dan *exhaustive hyperparameter optimization* dapat menghasilkan model klasifikasi yang tidak hanya akurat, tetapi juga efisien dan terinterpretasi. Dari perspektif metodologis, studi ini menawarkan kerangka *pipeline* yang terdokumentasi dan *reproducible*, mengatasi masalah umum dalam penelitian AI medis di mana konfigurasi *hyperparameter* dan protokol *preprocessing* sering kali tidak dilaporkan secara transparan [21], [23]. Secara praktis, temuan ini mendukung pergeseran paradigma dari sekadar mengejar akurasi tertinggi menuju pengembangan sistem pendukung keputusan klinis (CDSS) yang menekankan keseimbangan antara performa prediktif, kecepatan inferensi, dan transparansi fitur [1], [6]. Dengan membuktikan bahwa KNN teroptimasi dapat mencapai sensitivitas dan spesifisitas yang memadai untuk skrining awal, penelitian ini membuka peluang adopsi model berbasis instans yang lebih ringan secara komputasi di fasilitas kesehatan dengan infrastruktur terbatas, sejalan dengan rekomendasi Chen et al., dan Romeo et al., mengenai demokratisasi AI dalam hepatologi [3], [6].

Meskipun hasil penelitian menjanjikan, beberapa keterbatasan metodologis perlu diakui secara kritis. Pertama, ketidakseimbangan kelas (71,4% vs 28,6%) berpotensi memperkenalkan bias prediktif terhadap kelas mayoritas. Meskipun penggunaan metrik F1-Score dan *stratified sampling* telah memitigasi dampak ini, teknik *resampling* seperti SMOTE atau ADASYN belum diimplementasikan, yang dapat mempengaruhi kalibrasi probabilitas model [36], [37]. Kedua, ukuran dataset yang relatif kecil (n=583) membatasi kapasitas generalisasi ke populasi dengan variasi geografis, genetik, atau komorbiditas yang lebih beragam [5], [39]. Ketiga, evaluasi dilakukan menggunakan skema *hold-out* tunggal tanpa validasi eksternal pada kohort independen atau dataset multi-pusat, sehingga *robustness* model dalam skenario dunia nyata masih perlu diverifikasi [2], [6]. Ke depan, penelitian lanjutan disarankan untuk: (1) mengintegrasikan teknik penanganan *class imbalance* dan validasi silang *nested* untuk memastikan estimasi performa yang tidak bias [30], [33]; (2) melakukan uji coba prospektif pada dataset klinis *real-time* dari rumah sakit rujukan untuk menguji validitas ekologis model [3], [38]; dan (3) mengeksplorasi pendekatan *explainable AI* (XAI) seperti SHAP atau LIME untuk

memetakan kontribusi non-linear setiap fitur terhadap prediksi individu, sehingga meningkatkan kepercayaan klinis terhadap sistem otomatis [12], [22]. Dengan penyempurnaan metodologis tersebut, kerangka kerja yang diusulkan dalam studi ini memiliki potensi nyata untuk diadaptasi sebagai modul skrining non-invasif yang *scalable* dalam ekosistem kesehatan digital.

4. Simpulan

Penelitian ini berhasil mengevaluasi dan membandingkan kinerja algoritma *Support Vector Machine* (SVM) dan *K-Nearest Neighbor* (KNN) yang dioptimalkan melalui GridSearchCV dengan seleksi fitur berbasis *Chi-Square* pada klasifikasi penyakit hati. Hasil eksperimen secara konsisten menunjukkan bahwa integrasi *Chi-Square* dan GridSearchCV meningkatkan performa kedua model dibandingkan kondisi baseline, dengan reduksi dimensi dari 12 menjadi 7 fitur terpilih yang berhasil mempertahankan akurasi sekaligus meningkatkan efisiensi komputasi sebesar 28%. Model KNN teroptimasi (n_neighbors=9, metric=manhattan, weights=distance) mengungguli SVM teroptimasi (C=10, kernel=rbf) pada seluruh metrik evaluasi, mencapai akurasi 81,9%, presisi 83,5%, recall 81,9%, dan F1-Score 81,5%. Temuan ini mengonfirmasi bahwa kombinasi seleksi fitur statistik dan tuning hyperparameter sistematis merupakan pendekatan yang efektif untuk membangun model klasifikasi medis yang akurat, stabil, dan layak diintegrasikan ke dalam sistem pendukung keputusan klinis awal.

Secara praktis, ketujuh fitur yang teridentifikasi oleh Chi-Square merepresentasikan parameter biokimia dan klinis yang relevan secara medis, sehingga model tidak hanya unggul secara komputasi tetapi juga memiliki interpretabilitas tinggi bagi tenaga kesehatan. Namun, penelitian ini masih dibatasi oleh ketidakseimbangan distribusi kelas dan ukuran dataset yang relatif kecil, yang dapat membatasi generalisasi temuan ke populasi pasien yang lebih beragam secara demografis maupun geografis. Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan teknik penanganan class imbalance seperti SMOTE atau ADASYN, melakukan validasi eksternal pada kohort multi-pusat yang lebih besar, serta mengeksplorasi perbandingan komprehensif dengan algoritma ensemble modern seperti XGBoost atau arsitektur deep learning. Dengan pengembangan metodologis lebih lanjut, kerangka pipeline yang diusulkan dalam studi ini memiliki potensi signifikan untuk diadopsi sebagai alat skrining non-invasif yang efisien dalam deteksi dini dan manajemen penyakit hati.

Daftar Referensi

[1] C. Baciu, C. Xu, M. Alim, K. Prayitno, and M. Bhat, “Artificial intelligence applied to omics data in liver diseases: Enhancing clinical predictions,” *Front. Artif. Intell.*, vol. 5, pp. 1–10, Nov. 2022, doi: 10.3389/frai.2022.1050439.

[2] P. Lei *et al.*, “Radiobioinformatics: A Novel Bridge Between Basic Research and Clinical Practice for Clinical Decision Support in Diffuse Liver Diseases,” *Iradiology*, vol. 1, no. 2, pp. 167–189, 2023, doi: 10.1002/ird3.24.

[3] M. Romeo *et al.*, “Clinical Applications of Artificial Intelligence (AI) in Human Cancer: Is It Time to Update the Diagnostic and Predictive Models in Managing Hepatocellular Carcinoma (HCC)?,” *Diagnostics*, vol. 15, no. 3, p. 252, 2025, doi: 10.3390/diagnostics15030252.

[4] S. Singh, S. Hoque, A. Zekry, and A. Sowmya, “Radiological Diagnosis of Chronic Liver Disease and Hepatocellular Carcinoma: A Review,” *J. Med. Syst.*, vol. 47, no. 1, p. 73, Jul. 2023, doi: 10.1007/s10916-023-01968-7.

[5] A. Mansur *et al.*, “The Role of Artificial Intelligence in the Detection and Implementation of Biomarkers for Hepatocellular Carcinoma: Outlook and Opportunities,” *Cancers (Basel).*, vol. 15, no. 11, p. 2928, 2023, doi: 10.3390/cancers15112928.

[6] Z.-L. Chen, C. Wang, and F. Wang, “Revolutionizing gastroenterology and hepatology with artificial intelligence: From precision diagnosis to equitable healthcare through interdisciplinary practice,” *World J. Gastroenterol.*, vol. 31, no. 24, pp. 1–25, Jun. 2025, doi: 10.3748/wjg.v31.i24.108021.

[7] M. A. Abdel-Fattah, N. A. Othman, and N. Goher, “Predicting Chronic Kidney Disease Using Hybrid Machine Learning Based on Apache Spark,” *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–12, 2022, doi: 10.1155/2022/9898831.

[8] D. Swain *et al.*, “Differential Diagnosis of Erythemato-Squamous Diseases Using a Hybrid Ensemble Machine Learning Technique,” *Intelligent Decision Technologies*, vol. 18, no. 2, pp. 1495–1510, 2024, doi: 10.3233/idt-230779.

[9] R. T. E. Putri, J. Zeniarja, S. T. Winarno, A. N. Cahyani, and A. A. Maulani, "GridSearch and Data Splitting for Effectiveness Heart Disease Classification," *Sinkron*, vol. 9, no. 1, pp. 317–331, 2024, doi: 10.33395/sinkron.v9i1.13198.
[10] A. Rasool, C. Bunterngchit, T. Luo, Md. R. Islam, Q. Qu, and Q. Jiang, "Improved Machine Learning-Based Predictive Models for Breast Cancer Diagnosis," *Int. J. Environ. Res. Public Health*, vol. 19, no. 6, p. 3211, 2022, doi: 10.3390/ijerph19063211.
[11] F. Demir, K. Siddique, M. Alswaitti, K. Demir, and A. Şengür, "A Simple and Effective Approach Based on a Multi-Level Feature Selection for Automated Parkinson's Disease Detection," *J. Pers. Med.*, vol. 12, no. 1, p. 55, 2022, doi: 10.3390/jpm12010055.
[12] L. Kelebercová, M. Munk, and F. Forgáč, "Could You Understand Me? The Relationship Among Method Complexity, Preprocessing Complexity, Interpretability, and Accuracy," *Mathematics*, vol. 11, no. 13, p. 2922, 2023, doi: 10.3390/math11132922.
[13] T. A. Assegie, "Support vector machine and k-nearest neighbor based liver disease classification model," *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 3, no. 1, pp. 9–14, 2021, doi: 10.35882/ijeeemi.v3i1.2.
[14] I. B. S. Mahendra, T. Widiharih, F. A. Nugroho, and P. S. Sasongko, "Implementation of feature selection chi-square to improve the accuracy of the classification model using the random forest algorithm on coronary artery disease," *Journal of Applied Intelligent Systems*, vol. 9, no. 1, pp. 1–7, 2025, doi: 10.62411/jais.v9i1.7858.
[15] F. Zamachsari, G. V Saragih, and W. Gata, "Analisis sentimen pemindahan ibu kota negara dengan feature selection algoritma naive Bayes dan support vector machine," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 3, pp. 504–512, 2020, doi: 10.29207/resti.v4i3.1942.
[16] S. W. Iriananda, R. W. Budiawan, A. Y. Rahman, and I. Istiadi, "Optimasi klasifikasi sentimen komentar pengguna game bergerak menggunakan SVM, grid search dan kombinasi n-gram," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 743–752, 2024, doi: 10.25126/jtiik.1148244.
[17] K. N. Berawi, E. R. Susanto, A. Wantoro, M. N. D. Satria, H. Busman, and A. Wibowo, "Combination of XGBoost – Grid Search with SVM for Diabetes Diagnostics," in *2023 International Conference on Networking, Electrical Engineering, Computer Science, and Technology (IConnect)*, 2023, pp. 224–229. doi: 10.1109/iconnect56593.2023.10327016.
[18] I. J. Jebadurai, G. J. L. Paulraj, J. Jebadurai, and S. Silas, "Experimental Analysis of Filtering-Based Feature Selection Techniques for Fetal Health Classification," *Serbian Journal of Electrical Engineering*, vol. 19, no. 2, pp. 207–224, 2022, doi: 10.2298/sjee2202207j.
[19] H. M. Kim *et al.*, "Machine learning-based prediction model for late recurrence after surgery in patients with renal cell carcinoma," *BMC Med. Inform. Decis. Mak.*, vol. 22, no. 1, p. 241, Sep. 2022, doi: 10.1186/s12911-022-01964-w.
[20] C. Tuena, C. Pupillo, C. Stramba-Badiale, M. Stramba-Badiale, and G. Riva, "Predictive power of gait and gait-related cognitive measures in amnestic mild cognitive impairment: a machine learning analysis," *Front. Hum. Neurosci.*, vol. 17, pp. 1–11, Jan. 2024, doi: 10.3389/fnhum.2023.1328713.
[21] A. Zhang, L. Xing, J. Zou, and J. C. Wu, "Shifting machine learning for healthcare from development to deployment and from models to data," *Nature Biomedical Engineering 2022 6:12*, vol. 6, no. 12, pp. 1330–1345, Jul. 2022, doi: 10.1038/s41551-022-00898-y.
[22] A. Apicella, F. Isgrò, and R. Prevete, "Don't push the button! Exploring data leakage risks in machine learning and transfer learning," *Artificial Intelligence Review 2025 58:11*, vol. 58, no. 11, pp. 339-, Aug. 2025, doi: 10.1007/S10462-025-11326-3.
[23] A. Thakur, T. Zhu, V. Abrol, J. Armstrong, Y. Wang, and D. A. Clifton, "Data encoding for healthcare data democratization and information leakage prevention," *Nature Communications 2024 15:1*, vol. 15, no. 1, pp. 1582-, Feb. 2024, doi: 10.1038/s41467-024-45777-z.
[24] A. U. Rehman *et al.*, "A Machine Learning-Based Framework for Accurate and Early Diagnosis of Liver Diseases: A Comprehensive Study on Feature Selection, Data Imbalance, and Algorithmic Performance," *International Journal of Intelligent Systems*, vol. 2024, no. 1, p. 6111312, Jan. 2024, doi: 10.1155/2024/6111312.
[25] R. Amin, R. Yasmin, S. Ruhi, M. H. Rahman, and M. S. Reza, "Prediction of chronic liver disease patients using integrated projection based statistical feature extraction with

machine learning algorithms," *Inform. Med. Unlocked*, vol. 36, p. 101155, Jan. 2023, doi: 10.1016/J.IMU.2022.101155.

[26] W. Aprilliandhika and F. F. Abdulloh, "Comparison of K-Nearest Neighbor and Support Vector Machine Algorithm Optimization With Grid Search CV on Stroke Prediction," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 991–1000, Jul. 2024, doi: 10.52436/1.JUTIF.2024.5.4.1951.

[27] S. Mujiyono, U. P. Sanjaya, I. S. Wibisono, and H. Setyowati, "Prediksi Fluktuasi Berat Badan Berdasarkan Pola Hidup Menggunakan Model XGBoost dan Deep Learning," *Jurnal Algoritma*, vol. 22, no. 1, pp. 221–233, May 2025, doi: 10.33364/ALGORITMA/V.22-1.2253.

[28] S. Alelyani, "Stable bagging feature selection on medical data," *Journal of Big Data 2021 8:1*, vol. 8, no. 1, pp. 11-, Jan. 2021, doi: 10.1186/S40537-020-00385-8.

[29] M. U. Salur, "Balancing and feature selection strategy for Parkinson's disease classification: An improved sequential feature selection approach," *Biomed. Signal Process. Control*, vol. 113, p. 109061, Mar. 2026, doi: 10.1016/J.BSPC.2025.109061.

[30] S. Dhanka, A. Sharma, A. Kumar, S. Maini, and H. Vundavilli, "Advancements in Hybrid Machine Learning Models for Biomedical Disease Classification Using Integration of Hyperparameter-Tuning and Feature Selection Methodologies: A Comprehensive Review," *Archives of Computational Methods in Engineering 2025 33:1*, vol. 33, no. 1, pp. 289–324, Jun. 2025, doi: 10.1007/S11831-025-10309-5.

[31] C. Pavithra, M. Saradha, and V. Muthukumaran, "A Nonlinear SVM in Cardiovascular Diagnostics Disentangling Myocardial Infarction with 2D Data Analysis," *Computer Vision in Healthcare Prediction, Detection and Diagnosis*, pp. 141–155, Jan. 2026, doi: 10.1201/9781003541011-10/NONLINEAR-SVM-CARDIOVASCULAR-DIAGNOSTICS-PAVITHRA-SARADHA-MUTHUKUMARAN.

[32] Y. Wang, "Linear Dimensionality Reduction Methods for Analyzing Structured Biomedical Data: Existing Research and Future Opportunities," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 17, no. 3, p. e70045, Sep. 2025, doi: 10.1002/WICS.70045;PAGE:STRING:ARTICLE/CHAPTER.

[33] Q. A. Hidayaturrohman and E. Hanada, "A Comparative Analysis of Hyper-Parameter Optimization Methods for Predicting Heart Failure Outcomes," *Applied Sciences 2025, Vol. 15, Page 3393*, vol. 15, no. 6, p. 3393, Mar. 2025, doi: 10.3390/APP15063393.

[34] M. Jannah and A. A. Nababan, "Optimization of KNN Classification for ECG Data Analysis: Comparative Study of Model Performance Using the Hyperparameter Tuning and Cross-Validation," *International Journal of Electronics and Communications Systems*, vol. 5, no. 1, pp. 105–116, Jun. 2025, doi: 10.24042/IJECS.V5I1.28467.

[35] E. Hokijuliandy, H. Napitupulu, and Firdaniza, "Application of SVM and Chi-Square Feature Selection for Sentiment Analysis of Indonesia's National Health Insurance Mobile Application," *Mathematics*, vol. 11, no. 17, p. 3765, Sep. 2023, doi: 10.3390/math11173765.

[36] M. Salmi, D. Atif, D. Oliva, A. Abraham, and S. Ventura, "Handling imbalanced medical datasets: review of a decade of research," *Artificial Intelligence Review 2024 57:10*, vol. 57, no. 10, pp. 273-, Sep. 2024, doi: 10.1007/S10462-024-10884-2.

[37] A. Tiwari, S. Mishra, and S. Tiwari, "AI/ML: Developing Algorithms for Handling Imbalanced Datasets in Classification Tasks," *Advances in Science, Technology and Innovation*, vol. Part F703, pp. 17–29, 2025, doi: 10.1007/978-3-031-82733-4_2/SAVE-RESEARCH.

[38] R. Strandberg, P. Jepsen, and H. Hagström, "Developing and validating clinical prediction models in hepatology – An overview for clinicians," *J. Hepatol.*, vol. 81, no. 1, pp. 149–162, Jul. 2024, doi: 10.1016/J.JHEP.2024.03.030.

[39] M. S. Tudor *et al.*, "Evolutive Models, Algorithms and Predictive Parameters for the Progression of Hepatic Steatosis," *Metabolites 2024, Vol. 14, Page 198*, vol. 14, no. 4, p. 198, Apr. 2024, doi: 10.3390/METABO14040198.