Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

Analisis Komparatif Model Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, dan K-Nearest Neighbor untuk Klasifikasi Penyakit Batu Empedu Menggunakan Machine Learning

Ahda Rindang Al Amin, Budy Satria*

Fakultas Teknologi Informasi, Departemen Informatika, Universitas Andalas, Padang, Indonesia
Email: [1]ahdaalamin2506@gmail.com, [2,*]budy.satria@it.unand.ac.id
Email Korespondensi: budy.satria@it.unand.ac.id

Abstrak−Penyakit batu empedu merupakan kondisi medis yang cukup umum dan sering kali tidak menunjukkan gejala hingga mencapai tahap komplikasi. Prediksi dini terhadap penyakit ini dapat membantu meningkatkan kualitas hidup pasien melalui intervensi yang lebih cepat. Penelitian ini bertujuan untuk membandingkan performa lima algoritma klasifikasi dalam memprediksi penderita batu empedu berdasarkan data klinis, yaitu Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (K-NN). Dataset yang digunakan berasal dari UCI Machine Learning Repository dan berisi data klinis dari 319 pasien, dengan 38 fitur numerik. Penanganan outlier dilakukan menggunakan metode Winsorization dan RobustScaler. Setiap model dioptimalkan menggunakan GridSearchCV dan dievaluasi berdasarkan metrik akurasi, presisi, recall, F1-score, dan ROC AUC. Hasil menunjukkan bahwa algoritma SVM dan Logistic Regression memberikan hasil terbaik dengan akurasi sebesar 85,9% dan nilai AUC di atas 0,90. Berdasarkan temuan ini, Logistic Regression dan SVM direkomendasikan sebagai model klasifikasi yang paling efektif untuk prediksi penyakit batu empedu berbasis data klinis.

Kata Kunci: Batu Empedu; Klasifikasi; Machine Learning; Logistic Regression; Decision Tree; RF; SVM; KNN

Abstract−Gallstone disease is a common medical condition that often presents without symptoms until complications occur. Early prediction of this disease can improve patient outcomes through timely intervention. This study aims to compare the performance of five classification algorithms in predicting gallstone disease using clinical data: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (K-NN). The dataset used was obtained from the UCI Machine Learning Repository and contains clinical data from 319 patients, comprising 38 numerical features. Outlier handling was conducted using Winsorized Transformation and RobustScaler. Each model was optimized using GridSearchCV and evaluated using accuracy, precision, recall, F1-score, and ROC AUC metrics. The results show that SVM and Logistic Regression achieved the best performance, each with 85.9% accuracy and an AUC above 0.90. Based on these findings, Logistic Regression and SVM are recommended as the most effective classification models for gallstone disease prediction using clinical data.

Keywords: Gallstone; Classification; Machine Learning; Logistic Regression; Decision Tree; RF; SVM; KNN

1. PENDAHULUAN

Batu empedu atau yang secara ilmiah disebut Cholelithis merupakan penyakit yang dapat ditemukan di dalam kandungan empedu atau di dalam saluran empedu atau bahkan pada keduanya [1]. Penyakit batu empedu adalah istilah yang digunakan untuk merujuk pada keberadaan batu yang biasanya terbentuk di dalam kantung empedu atau di saluran empedu dan saluran hati. Batu empedu adalah endapan yang terbentuk di kantong empedu, sering kali disebabkan oleh kelebihan kolesterol atau bilirubin [2]. Lebih dari 90 % batu empedu adalah batu kolesterol (komposisi kolesterol >50 %) atau bentuk campuran (20-50 % unsur kolesterol) dan sisanya 10 % adalah batu pigmen (unsur kalsium dominan dan kolesterol 20%). Menurut gambaran makroskopik dan komposisi kimianya, batu saluran empedu dapat diklasifikasikan menjadi kategori mayor, yaitu batu pigmen cokelat atau batu kalsium bilirubinate yang mengandung kalsium bilirubinate sebagai komponen utama, dan batu pigmen hitam yang kaya akan residu hitam tak terekstraksi. Hampir 50% penderita batu empedu tidak merasakan gejala apa-apa, 30% merasakan gejala nyeri, dan 20% berkembang menjadi komplikasi. Gejala penyakit batu empedu dapat berkisar dari ringan dan tidak spesifik hingga nyeri hebat atau komplikasi. Gejala-gejala ini dapat mencakup nyeri berat di daerah perut atas, sakit kuning, demam (jika timbul komplikasi), dan muntah-muntah. Faktor risiko batu empedu dikenal dengan singkatan 4F, yaitu Forty, Female, Fat, Family. Artinya, batu empedu lebih umum pada mereka yang berusia di atas 40 tahun, wanita, yang gemuk, dan yang punya riwayat keluarga terkena batu empedu. Faktor risiko batu empedu jarang sekali menyerang orang berusia 25 tahun ke bawah.

Selain itu, wanita lebih banyak terkena batu empedu dibandingkan dengan pria. Wanita memiliki prevalensi lebih tinggi dibandingkan dengan pria. Ditemukan pasien wanita memiliki persentase sebesar 61,8% yang disebabkan oleh hormon estrogen terhadap terjadinya peningkatan ekskresi kolesterol [3]. Orang dengan usia >40 tahun memiliki kecenderungan terkena kolelitiasis dibandingkan dengan usia yang lebih muda dikarenakan peningkatan sekresi kolesterol ke dalam empedu sesuai dengan bertambahnya usia [4]. Dalam era perkembangan teknologi yang pesat, machine learning menjadi salah satu pendekatan penting dalam berbagai bidang, termasuk kesehatan. Machine learning menawarkan solusi untuk menganalisis data secaramandiri tanpa pengawasan, sehingga mempermudah diagnosis penyakit dan pengambilan keputusan klinis [5]. Seiring dengan perkembangan kecerdasan buatan, pembelajaran mesin telah menjadi salah satu pendekatan yang menjanjikan untuk memprediksi risiko penyakit [6]. Namun, secara spesifik, studi yang membahas prediksi penyakit batu empedu dengan pendekatan data mining masih tergolong terbatas, Salah satu penyebab utamanya adalah keterbatasan data klinis yang tersedia. Salah satu penelitian yang pernah dilakukan [7], Dalam penelitian tersebut, atribut yang digunakan dalam proses klasifikasi mencakup berat badan, faktor genetik, riwayat infeksi, dan diabetes.

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

Dengan menerapkan algoritma C4.5, diperoleh pohon keputusan yang menunjukkan bahwa faktor diabetes dan genetik merupakan prediktor utama dari penyakit batu empedu. Hasil klasifikasinya menghasilkan beberapa aturan, seperti: jika seseorang menderita diabetes, maka kemungkinan besar ia mengidap batu empedu; jika tidak menderita diabetes, maka kecenderungannya tidak mengidap batu empedu; namun, apabila tidak menderita diabetes tetapi memiliki faktor genetik, maka ia tetap berisiko mengidap batu empedu; sebaliknya, jika tidak menderita diabetes dan tidak memiliki riwayat genetik, maka kecil kemungkinan mengidap batu empedu.

Permasalahan pada penelitian ini adalah penyakit batu empedu, yang merupakan salah satu gangguan sistem pencernaan yang memerlukan diagnosis yang tepat untuk mendukung penanganan medis secara dini. Penelitian terkait klasifikasi penderita penyakit batu empedu belum banyak dilakukan. Perkembangan machine learning telah menawarkan berbagai algoritma klasifikasi yang mampu membantu proses prediksi penyakit berdasarkan data klinis. Namun, setiap algoritma memiliki karakteristik dan performa yang berbeda bergantung pada kondisi dataset, seperti keberadaan *outlier*, distribusi data, serta hubungan antarfitur. Selain itu, belum banyak penelitian yang membandingkan secara komprehensif performa *Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), dan K-Nearest Neighbor (K-NN)* pada dataset klinis penyakit batu empedu dengan pendekatan penanganan *outlier* menggunakan *Winsorized Transformation* dan *RobustScaler*.

Penelitian terdahulu menunjukkan bahwa berbagai algoritma machine learning dan pendekatan probabilistik telah digunakan untuk mendukung diagnosis dan klasifikasi penyakit, termasuk batu empedu, diabetes, dan penyakit jantung. Penelitian mengenai prediksi status komorbiditas pada pasien batu empedu menggunakan algoritma Gradient Boosted Tree (GBT) dan menghasilkan akurasi sebesar 82,29%, classification error 17,71%, weighted mean precision 89,63%, serta weighted mean recall 72,58% [8]. Meskipun demikian, penelitian tersebut belum melakukan optimasi hyperparameter secara komprehensif dan hanya menggunakan satu model. Selanjutnya, penelitian diagnosis tingkat kepastian penyakit batu empedu dengan Teorema Bayes memperoleh tingkat kepastian sebesar 69,02%, tetapi pendekatan yang digunakan masih terbatas pada satu model [1]. Pada klasifikasi penyakit diabetes, perbandingan Random Forest dan Support Vector Machine menunjukkan bahwa Random Forest dengan pembagian data 80%:20% memberikan kinerja terbaik, dengan akurasi 0,98, presisi 0,96, recall 1, specificity 0,95, dan F1-score 0,98 [9]. Namun, penelitian tersebut belum mengevaluasi teknik penanganan outlier, seperti Winsorized Transformation atau RobustScaler. Sementara itu, klasifikasi penyakit jantung menggunakan K-Nearest Neighbor memperoleh accuracy 92%, precision 90%, dan recall 92% [10]. Berdasarkan penelitian terdahulu, masih terdapat peluang untuk mengembangkan penelitian melalui perbandingan beberapa algoritma, optimasi hyperparameter, serta penerapan strategi prapemrosesan data yang lebih robust untuk meningkatkan kinerja klasifikasi penyakit.

Berdasarkan analisis kesenjangan penelitian terdahulu pada Tabel 1, masih terdapat variasi dalam strategi prapemrosesan data, khususnya penanganan *outlier* dan keterbatasan dalam jumlah algoritma yang dibandingkan serta proses optimasi *hyperparameter*. Oleh karena itu, penelitian ini mengusulkan evaluasi komprehensif terhadap lima algoritma machine learning dengan menerapkan penanganan *outlier* menggunakan Winsorized Transformation dan RobustScaler, optimasi *hyperparameter* melalui GridSearchCV, serta evaluasi menggunakan akurasi, presisi, *recall*, F1-score, dan ROC AUC.

Oleh karena itu, penelitian ini bertujuan untuk melakukan evaluasi komparatif kelima algoritma tersebut untuk menentukan model klasifikasi yang memiliki performa terbaik berdasarkan metrik akurasi, presisi, *recall*, F1-score, dan ROC AUC. Selain itu, penelitian ini juga bermaksud untuk membandingkan beberapa algoritma klasifikasi, yaitu *Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (K-NN)*, untuk memprediksi penderita penyakit batu empedu. Kelima algoritma ini dipilih karena masing-masing memiliki pendekatan yang berbeda dalam menangani data klasifikasi, sehingga memberikan perspektif yang lebih komprehensif terhadap performa prediksi penyakit batu empedu.

Logistic Regression (LR) sering dijadikan model baseline karena kesederhanaan dan kemudahan interpretasi, yang penting dalam konteks medis untuk memvalidasi temuan klinis sebelum menerapkan model yang lebih kompleks [11]. Logistic regression merupakan salah satu jenis regresi yang mengubungkan antara satu atau beberapa variabel independen (variabel bebas) dengan variabel dependen yang berupa kategori, biasanya 0 dan 1 [12]. *Decision Tree* merupakan algoritma Tree biasa dipakai untuk pengenalan pola statistik. Decision Tree terbuat dari tiga simpul yaitu leaf, lalu terdiri juga dari simpul root yang merupakan titik awal dari suatu decision tree, dan yang terakhir adalah simpul perantara yang berhubungan dengan suatu pengujian [13]. Penelitian ini juga menggunakan model *Random Forest* (RF) untuk mengurangi risiko overfitting yang sering terjadi pada pohon keputusan (decision tree) tunggal, sehingga meningkatkan keandalan prediksi baik pada tugas klasifikasi maupun regresi. Selain mampu mengatasi overfitting, Random Forest juga memiliki berbagai keunggulan, seperti fleksibilitas dalam menangani data yang heterogen, toleransi terhadap nilai yang hilang dan outlier, serta kemampuan untuk memberikan informasi mengenai tingkat kepentingan dari setiap fitur yang digunakan dalam proses pemodelan [14]. RF merupakan salah satu metode berbasis klasifikasi dan regresi di mana terdapat proses agregasi pohon keputusan. Metode ini dogunakan karena menghasilkan kesalahan yang lebih rendah, memberikan akurasi yang bagus dalam klasifikasi dan dapat menangani data pelatihan yang jumlahnya sangat besar [15]. Support Vector Machine (SVM) merupakan algoritma yang mampu menangani proses optimasi secara efisien melalui pendekatan komputasi numerik. Untuk memperoleh solusi yang optimal, SVM memecah permasalahan optimasi yang kompleks menjadi sejumlah submasalah yang lebih sederhana [16]. *Support Vector Machine* (SVM) merupakan suatu teknik untuk menemukan hyperplane yang bisa memisahkan dua set data dari dua kelas yang berbeda. SVM memiliki kelebihan diantaranya adalah dalam menentukan jarak menggunakan support vector sehingga

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

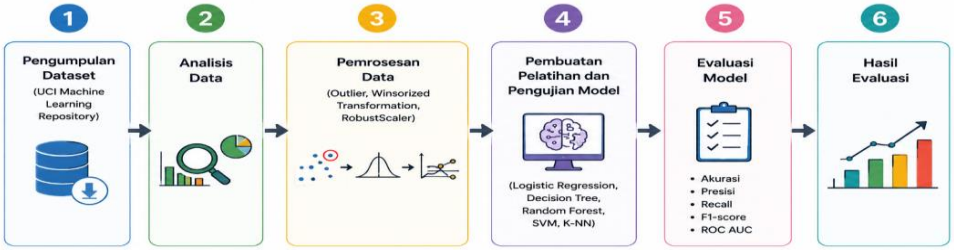
proses komputasi menjadi cepat [17]. Sedangkan KNN merupakan metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [18]. Pemilihan lima algoritma tersebut diharapkan mampu memberikan gambaran yang menyeluruh mengenai efektivitas masing-masing model dalam mengklasifikasikan penderita batu empedu. Evaluasi dilakukan dengan membandingkan performa masing-masing model berdasarkan metrik akurasi, presisi, recall, F1-score, dan ROC-AUC terhadap dataset yang telah diproses.

Penelitian ini memberikan tiga kontribusi utama. Pertama, melakukan evaluasi komparatif lima algoritma machine learning pada dataset klinis penyakit batu empedu menggunakan skenario eksperimen yang seragam sehingga menghasilkan perbandingan performa yang lebih objektif. Kedua, menerapkan kombinasi Winsorized Transformation dan RobustScaler sebagai strategi penanganan outlier yang disesuaikan dengan karakteristik data sebelum proses pemodelan. Ketiga, mengoptimalkan parameter setiap algoritma melalui GridSearchCV dan mengevaluasi performanya menggunakan metrik akurasi, presisi, recall, F1-score, dan ROC AUC untuk memperoleh model klasifikasi yang lebih baik.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan pendekatan eksperimen berbasis machine learning dengan alur diawali dengan pengumpulan dataset yang diperoleh dari UCI Machine Learning Repository sebagai sumber data penelitian. Selanjutnya dilakukan analisis data untuk memahami karakteristik dataset, termasuk distribusi data, jenis atribut, serta hubungan antarvariabel. Pada tahap pemrosesan data, dilakukan identifikasi data outlier untuk mendeteksi nilai-nilai ekstrem, kemudian dilakukan normalisasi menggunakan Winsorized Transformation guna mengurangi pengaruh outlier dan RobustScaler untuk menyelaraskan skala data sehingga lebih sesuai untuk proses pemodelan. Setelah data diproses, dilakukan pembuatan model dengan membagi dataset menjadi data pelatihan (training set) dan data pengujian (testing set), kemudian menerapkan lima algoritma machine learning, yaitu *Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (K-NN).* Selanjutnya, setiap model dievaluasi menggunakan metrik akurasi, presisi, recall, F1-score, serta Receiver Operating Characteristic–Area Under the Curve (ROC AUC) untuk mengukur kemampuan klasifikasi masing-masing algoritma. Tahap akhir penelitian menghasilkan hasil evaluasi berupa perbandingan performa kelima algoritma sehingga dapat diidentifikasi model dengan kinerja terbaik yang paling efektif untuk klasifikasi penderita batu empedu. Alur tahapan penelitian secara lengkap disajikan pada Gambar 1.



Gambar 1. Tahapan Penelitian yang diusulkan

Pada Gambar 1 dapat dijelaskan sebagai berikut:

1. Pengumpulan Data

Penelitian ini menggunakan dataset Gallstone yang diunduh dari UCI Machine Learning Repository bersumber dari dataset yang tersedia di platform Kaggle https://www.kaggle.com/datasets/xixama/gallstone-dataset-uci. Dataset ini berasal dari Klinik Rawat Jalan Penyakit Dalam di Ankara VM Medical Park Hospital, Turki, dan mencakup data klinis dari 319 individu yang dikumpulkan selama periode Juni 2022 hingga Juni 2023. Dari keseluruhan data tersebut, sebanyak 161 individu terdiagnosis menderita penyakit batu empedu, sementara sisanya tidak, sehingga menjadikan dataset ini seimbang antara kelas positif dan negatif. Dataset ini terdiri dari 38 fitur yang mencerminkan informasi klinis dan biologis pasien. Dengan kelengkapan dan kualitas data yang tinggi, dataset ini menyediakan fondasi yang kuat untuk eksplorasi model klasifikasi dalam memprediksi keberadaan penyakit batu empedu berbasis fitur klinis non-imaging.

2. Analisis Data

Tahap awal dalam penelitian ini adalah melakukan analisis terhadap struktur dan isi dataset. Analisis ini bertujuan untuk memahami karakteristik data, mengidentifikasi nilai ekstrem (outlier), serta mengevaluasi distribusi nilai dari masing-masing fitur. Meskipun dataset telah dinyatakan lengkap tanpa nilai kosong dan memiliki proporsi yang seimbang antara kelas penderita dan non-penderita batu empedu, pemeriksaan lebih lanjut menunjukkan adanya beberapa nilai outlier yang dapat memengaruhi performa model. Oleh karena itu, perlu dilakukan penanganan terhadap outlier sebelum data digunakan dalam proses pelatihan.

3. Pra-Pemrosesan Data

Pada tahap pra-pemrosesan data, dilakukan beberapa langkah penting untuk menyiapkan data agar optimal bagi proses klasifikasi. Langkah pertama adalah mengidentifikasi nilai outlier menggunakan metode Z-score dengan threshold 3. Karena jumlah data yang terbilang sedikit, pendekatan dengan melakukan penghapusan data outlier dinilai kurang tepat

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

karena dapat menyebabkan hilangnya informasi yang relevan. Sebagai alternatif, digunakan pendekatan normalisasi yang mampu meredam pengaruh outlier tanpa menghilangkan data. Dua metode utama yang digunakan adalah *Winsorized Transformation* dan *RobustScaler*. *Winsorized Transformation* adalah metode transformasi nilai ekstrem dengan membatasi nilai-nilai outlier pada ambang batas tertentu (biasanya persentil ke-5 dan ke-95). Artinya, nilai yang lebih rendah dari persentil ke-5 akan diubah menjadi nilai persentil ke-5, dan nilai yang lebih tinggi dari persentil ke-95 akan dipangkas menjadi nilai persentil ke-95. Tujuannya adalah meredam pengaruh outlier tanpa menghilangkan baris data, sehingga distribusi tetap terjaga secara umum, namun tidak didistorsi oleh nilai ekstrem.

RobustScaler, di sisi lain, adalah metode standardisasi yang tidak sensitif terhadap outlier karena menggunakan median dan interquartile range (IQR) dalam proses transformasinya, bukan mean dan standard deviation seperti *StandardScaler*. Skala baru dari setiap fitur dihitung dengan rumus:

$$X_{scaled} = \frac{X - median}{IQR} \quad (1)$$

Metode ini sangat cocok digunakan ketika data memiliki distribusi yang tidak normal atau mengandung nilai *outlier*.

4. Pelatihan dan Pengujian Model

Setelah data diproses, dilakukan proses pembuatan model klasifikasi menggunakan lima algoritma berbeda, yaitu *Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (K-NN).* Masing-masing model dibangun menggunakan pustaka scikit-learn, dan dioptimalkan melalui proses tuning hyperparameter menggunakan teknik *Grid Search Cross Validation*. Dengan metode ini, parameter-parameter model dapat disesuaikan untuk menghasilkan kinerja terbaik. Model yang telah dibangun kemudian diuji pada data uji untuk menilai kemampuannya dalam melakukan klasifikasi terhadap data yang belum pernah dilihat sebelumnya.

5. Evaluasi Model

Evaluasi dilakukan untuk mengukur performa model secara menyeluruh. Metrik yang digunakan meliputi akurasi, presisi, recall, dan F1-score, yang masing-masing memberikan gambaran mengenai kemampuan model dalam mengklasifikasikan data secara benar. Selain itu, evaluasi juga dilengkapi dengan metrik ROC AUC (Receiver Operating Characteristic – Area Under Curve), yang mengukur kemampuan model dalam membedakan antara dua kelas secara menyeluruh pada berbagai ambang keputusan (threshold).

Visualisasi berupa confusion matrix juga digunakan untuk memperlihatkan jumlah prediksi benar dan salah secara eksplisit. Dari hasil evaluasi ini, dilakukan perbandingan antarmodel untuk menentukan algoritma mana yang paling efektif dalam klasifikasi penderita penyakit batu empedu.

6. Hasil Evaluasi

Tahap hasil evaluasi bertujuan untuk membandingkan kinerja seluruh algoritma machine learning yang telah dibangun berdasarkan nilai metrik evaluasi, yaitu akurasi, presisi, recall, F1-score, dan ROC AUC. Perbandingan tersebut digunakan untuk mengidentifikasi algoritma yang memberikan performa klasifikasi terbaik terhadap data penderita batu empedu. Hasil evaluasi selanjutnya dianalisis sebagai dasar dalam menentukan model yang paling efektif, memiliki kemampuan generalisasi yang baik, serta layak direkomendasikan untuk mendukung proses klasifikasi penderita batu empedu pada penelitian ini.

3. HASIL DAN PEMBAHASAN

3.1 Fitur dan Label Data

Data Gallstone merupakan data yang berisi informasi data klinis dan biologis pasien serta ada atau tidaknya pasien tersebut menderita batu empedu. Data ini mempunyai 38 fitur, 1 buah label, dan terdiri dari 319 record. Seluruh fitur yang terdapat pada data yang digunakan sudah berbentuk numerik, baik itu integer, biner, kontinu, maupun kategorikal dengan angka (0, 1, 2, dst.). Label data yang digunakan pada dataset terdapat pada kolom "*Gallstone Status*" yang bernilai 0 untuk penderita batu empedu, dan 1 untuk yang sehat. Data selanjutnya diolah melalui bahasa pemrograman Python dengan menggunakan Google Colab sebagai IDE (Integrated Development Environment). Setiap fitur dan label yang digunakan pada penelitian ini dapat dilihat pada Tabel 1 sebagai berikut.

Tabel 1. Variabel Data *Gallstone*

No	Nama Variabel	Keterangan	Tipe Data	Role
1	*Gallstone Status*	Penderita batu empedu (0), dan sehat(1)	Binary	Label
2	*Age*	Usia pasien	Integer	Fitur
3	*Gender*	Jenis kelamin pasien	Categorical	Fitur
4	*Comorbidity*	Jumlah penyakit penyerta	Categorical	Fitur
5	*Coronary Artery Disease (CAD)*	Penyakit jantung koroner	Binary	Fitur
6	*Hypothyroidism*	Penyakit gangguan pada tiroid	Binary	Fitur

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

No	Nama Variabel	Keterangan	Tipe Data	Role
7	*Hyperlipidemia*	Kondisi kadar lemak tinggi dalam darah	Binary	Fitur
8	*Diabetes Mellitus (DM)*	Kondisi gula darah tinggi	Binary	Fitur
9	*Height*	Tinggi badan	Integer	Fitur
10	*Weight*	Berat badan	Continuous	Fitur
11	*Body Mass Index (BMI)*	Rasio berat dan tinggi badan	Continuous	Fitur
12	*Total Body Water (TBW)*	Jumlah air dalam tubuh pasien	Continuous	Fitur
13	*Extracellular Water (ECW)*	Body water di luar sel-sel tubuh	Continuous	Fitur
14	*Intracelluler Water (ICW)*	Body water di dalam sel-selu tubuh	Continuous	Fitur
15	*Extracellular Fluid/Total Body Water (ECF/TBW)*	Rasio ECW dengan TBW	Continuous	Fitur
16	*Total Body Fat Ratio (TBFR)*	Persentase total lemak dalam tubuh	Continuous	Fitur
17	*Lean Mass (LM)*	Persentase massa tubuh bukan lemak	Continuous	Fitur
18	*Body Protein Content*	Persentase protein dalam tubuh	Continuous	Fitur
19	*Visceral Fat Rating (VFR)*	Skor jumlah lemak viseral	Integer	Fitur
20	*Bone Mass (BM)*	Berat total tulang	Continuous	Fitur
21	*Muscle Mass (MM)*	Jumlah total otot	Continuous	Fitur
22	*Obesity*	Persentase lemak tubuh berlebih	Continuous	Fitur
...	...	...	...	...
...	...	...	...	...
32	*Aspartat Aminotransferaz (ALT)*	Kadar enzim ALT (enzim hati)	Continuous	Fitur
33	*Alkaline Phosphatae (ALP)*	Kadar enzim ALP (enzim hati)	Continuous	Fitur
34	*Creatinine*	Produk limbah dari otot dan disaring oleh ginjal	Continuous	Fitur
35	*Glomerular Filtration Rate (GFR)*	Ukuran seberapa baik ginjal menyaring limbah dari darah	Continuous	Fitur
36	*C-Reactive Protein (CRP)*	Penanda peradangan dalam tubuh	Continuous	Fitur
37	*Hemoglobin (HGB)*	Protein dalam sel darah merah	Continuous	Fitur
38	*Vitamin D*	Kadar vitamin D	Continuous	Fitur

Selain itu juga terdapat informasi tambahan dari variabel data Gallstone, yakni deskripsi dari fitur-fitur dengan tipe data biner dan kategorikal pada Tabel 2 sebagai berikut:

Tabel 2. Deskripsi Fitur Tipe Data Binary dan Categorical

No	Nama Variabel	Keterangan
1	*Gallstone Status*	0 (No) , 1 (Yes)
2	*Gender*	0 (Male), 1 (Female)
3	*Comorbidity*	0 (No comorbidities present), 1 (One comorbid condition), 2 (Two comorbid conditions), 3 (Three or more comorbid conditions)
4	*Coronary Artery Disease (CAD)*	0 (*No*), 1 (*Yes*)
5	*Hypothyroidism*	0 (*No*), 1 (*Yes*)
6	*Hyperlipidemia*	0 (*No*), 1 (*Yes*)
7	*Diabetes Mellitus*	0 (*No*), 1 (*Yes*)
8	*Hepatic Fat Accumulation (HFA)*	0 (*No fat accumulation*),1 (*Grade 1 (mild)*), 2 (*Grade 2 (moderate)*), 3 (*Grade 3 (severe)*), 4 (*Grade 4 (very severe)*)

3.2 Eksplorasi Data

Sebagai langkah awal dalam memahami struktur dan karakteristik dataset, dilakukan eksplorasi data untuk melihat representasi awal dari data yang digunakan. Data diperiksa sebelum dibangunnya model, sehingga didapatkan wawasan maksimal dari dataset yang dimiliki [19]. Pertama, lima baris pertama dari dataset ditampilkan untuk memberi gambaran umum mengenai fitur-fitur yang tersedia, terlihat pada Gambar 2 sebagai berikut.

	Gallstone Status	Age	Gender	Comorbidity	Coronary Artery Disease (CAD)	Hypothyroidism	Hyperlipidemia	Diabetes Mellitus (DM)	Height	Weight	...
0	0	50	0	0	0	0	0	0	185	92.8	...
1	0	47	0	1	0	0	0	0	176	94.5	...
2	0	61	0	0	0	0	0	0	171	91.1	...
3	0	41	0	0	0	0	0	0	168	67.7	...
4	0	42	0	0	0	0	0	0	178	89.6	...

(a)

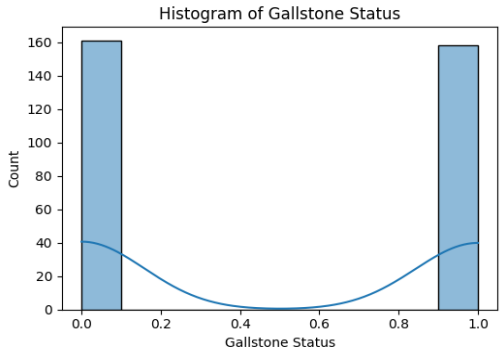
Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

High Density Lipoprotein (HDL)	Triglyceride	Aspartat Aminotransferaz (AST)	Alanin Aminotransferaz (ALT)	Alkaline Phosphatase (ALP)	Creatinine	Glomerular Filtration Rate (GFR)	C-Reactive Protein (CRP)	Hemoglobin (HGB)	Vitamin D
40.0	134.0	20.0	22.0	87.0	0.82	112.47	0.0	16.0	33.0
43.0	103.0	14.0	13.0	46.0	0.87	107.10	0.0	14.4	25.0
43.0	69.0	18.0	14.0	66.0	1.25	65.51	0.0	16.2	30.2
59.0	53.0	20.0	12.0	34.0	1.02	94.10	0.0	15.4	35.4
30.0	326.0	27.0	54.0	71.0	0.82	112.47	0.0	16.8	40.6

(b)

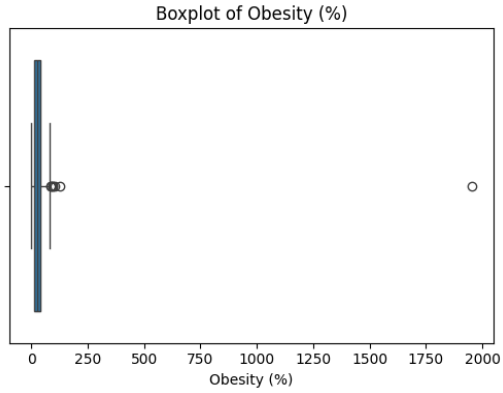
Gambar 2. Lima Baris Teratas Data Gallstone

Selanjutnya sebaran data masing-masing fitur ditampilkan menggunakan visualisasi histogram bisa dilihat Pada Gambar 3 sebagai berikut:



Gambar 3. Sebaran Data Label

Berdasarkan sebaran data, pada Gambar 3 menampilkan histogram distribusi kelas target atau label, yaitu penderita batu empedu (label 0) dan non-penderita (label 1). Dari grafik ini dapat dilihat bahwa jumlah data antara dua kelas relatif seimbang, yang merupakan kondisi ideal untuk membangun model klasifikasi. Selanjutnya, sebaran data tiap fitur juga ditampilkan dengan visualisasi data yang berbeda, yaitu *boxplot*. *Boxplot* menampilkan sebaran data menggunakan rentang interkuartil. Dengan ini bisa juga terlihat indikasi *outliers* dari masing-masing fitur. Salah satu fitur yang mengandung outlier yaitu Obesity dapat dilihat pada Gambar 4 sebagai berikut:



Gambar 4. Boxplot Fitur Obesity

Gambar 4 memperlihatkan boxplot dari salah satu fitur, yaitu Obesity, yang diambil sebagai contoh fitur yang mengandung nilai outlier. Terlihat adanya nilai yang berada jauh di luar rentang interkuartil, yang diindikasikan oleh titik-titik di luar whisker boxplot. Nilai-nilai outlier ini berpotensi memengaruhi performa model.

3.3 Outlier Detection

Setelah dilakukan eksplorasi awal, ditemukan bahwa beberapa fitur dalam dataset mengandung nilai-nilai outlier, yaitu nilai yang secara signifikan berbeda dari mayoritas distribusi data. Outlier seperti ini dapat memengaruhi performa model klasifikasi. Oleh karena itu, penanganan outlier menjadi langkah penting dalam tahap pra-pemrosesan data. Dalam penelitian ini, outlier diidentifikasi menggunakan metode Z-Score dengan ambang batas (threshold) sebesar 3, yang berarti setiap nilai dengan skor Z lebih dari ±3 dianggap sebagai outlier. Penanganan ini diterapkan hanya pada fitur numerik kontinu, sementara fitur dengan tipe biner atau kategorikal dikecualikan agar tidak terjadi distorsi nilai. Fitur yang diterapkan dan jumlah outlier pada masing-masing fitur dapat dilihat pada Tabel 3.

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

Tabel 3. Jumlah Outlier pada masing-masing fitur

No	Nama Variabel	Keterangan
1	*Body Mass Index (BMI)*	4 outliers
2	*Total Body Water (TBW)*	2 outliers
3	*Extracellular Water (ECW)*	1 outliers
4	*Intracellular Water (ICW)*	1 outliers
5	*Extracellular Fluid/Total Body Water (ECF/TBW)*	4 outliers
6	*Total Body Fat Ratio (TBFR) (%)*	0 outliers
7	*Lean Mass (LM) (%)*	0 outliers
8	*Body Protein Content (Protein) (%)*	3 outliers
9	*Visceral Fat Rating (VFR)*	3 outliers
10	*Bone Mass (BM)*	0 outliers
11	*Muscle Mass (MM)*	1 outliers
12	*Obesity (%)*	1 outliers
13	*Total Fat Content (TFC)*	Total Fat Content (TFC)
14	*Visceral Fat Area (VFA)*	Visceral Fat Area (VFA)
15	*Visceral Muscle Area (VMA) (Kg)*	Visceral Muscle Area (VMA) (Kg)
16	*Hepatic Fat Accumulation (HFA)*	Hepatic Fat Accumulation (HFA)
17	*Glucose*	Glucose
18	*Total Cholesterol (TC)*	Total Cholesterol (TC)
19	*Low Density Lipoprotein (LDL)*	Low Density Lipoprotein (LDL)
20	*High Density Lipoprotein (HDL)*	High Density Lipoprotein (HDL)
21	*Triglyceride*	Triglyceride
…	…	…
…	…	…
29	*Vitamin D*	2 outliers

Setelah dilakukan identifikasi terhadap nilai-nilai outlier berdasarkan Z-score dengan threshold 3, diketahui bahwa terdapat 49 baris data yang memiliki setidaknya satu fitur dengan nilai outlier. Jumlah ini tergolong cukup besar jika dibandingkan dengan total data yang tersedia, yaitu 319 baris. Menghapus baris-baris tersebut berisiko mengurangi variasi data yang penting dan dapat menyebabkan kehilangan informasi yang berharga dalam proses klasifikasi. Oleh karena itu, pendekatan alternatif dilakukan dengan normalisasi data, bukan penghapusan.

3.4 Normalisasi

Proses normalisasi dilakukan dengan mempertimbangkan karakteristik masing-masing model klasifikasi. Digunakan tiga pendekatan utama, yaitu Winsorized Transformation dan RobustScaler, serta gabungan dari keduanya. Dalam penelitian ini, pemilihan metode normalisasi dilakukan secara adaptif terhadap model. Untuk Logistic Regression, digunakan RobustScaler saja, karena model ini cukup sensitif terhadap distribusi variabel dan lebih stabil dengan skala yang tahan terhadap outlier. Untuk K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), dan Decision Tree, digunakan gabungan Winsorized dan RobustScaler, karena algoritma tersebut dipengaruhi baik oleh jarak antartitik maupun oleh kepekaan terhadap sebaran nilai ekstrem. Sementara itu, Random Forest menggunakan Winsorized Transformation saja, karena sebagai model ensemble berbasis pohon, Random Forest cukup tahan terhadap skala data namun tetap bisa terganggu oleh nilai ekstrem yang terlalu jauh.

3.5 Pembagian Data dan Pelatihan Model

Setelah data dinormalisasi, tahap berikutnya adalah membagi dataset menjadi data latih dan data uji. Proses ini dilakukan untuk mengevaluasi kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya. Pembagian dilakukan dengan rasio 80:20, di mana 80% data digunakan sebagai data latih dan 20% sisanya sebagai data uji. Untuk memastikan konsistensi hasil, pembagian data dilakukan dengan parameter random_state=42. Mengingat adanya tiga versi data hasil normalisasi, yaitu data dengan transformasi Winsorized, RobustScaler, dan gabungan Winsorized+RobustScaler, dilakukan tiga kali proses pembagian, masing-masing sesuai dengan metode normalisasi yang digunakan oleh model terkait. Variabel target (y) tetap sama pada ketiga pembagian data. Setelah data dibagi, dilakukan tahap inisialisasi dan pelatihan (training) terhadap lima algoritma klasifikasi, yaitu Logistic Regression, K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), Decision Tree, dan Random Forest. Model kemudian dilatih pada data latih masing-masing menggunakan fungsi *.fit()*.

3.6 Tuning Hyperparameter

Setelah model dasar (*baseline*) berhasil dibangun dan dilatih, tahap selanjutnya adalah melakukan *tuning hyperparameter*. Proses ini berguna untuk meningkatkan performa masing-masing algoritma klasifikasi. *Tuning hyperparameter* bertujuan untuk menemukan kombinasi parameter terbaik yang menghasilkan akurasi dan generalisasi model yang optimal terhadap data uji. Dalam penelitian ini, *tuning* dilakukan menggunakan metode *Grid Search Cross Validation* (GridSearchCV) dari pustaka scikit-learn untuk memilih parameter terbaik pada masing-masing algoritma sehingga ditemukan performa yang terbaik [20]. Metode ini melakukan pencarian secara menyeluruh terhadap kombinasi *hyperparameter* yang telah ditentukan sebelumnya, dengan evaluasi berbasis validasi silang (*cross-validation*). Setiap kombinasi parameter diuji,

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

dan hasil terbaik dipilih berdasarkan metrik akurasi rata-rata. Berikut adalah hasil terbaik (*best parameters*) dari tuning *hyperparameter* untuk masing-masing model yang bisa dilihat pada Tabel 4 sebagai berikut.

Tabel 4. Hasil Tuning Hyperparameter

Model	*Hyper-parameter*	Hasil Terbaik	Keterangan
Logistic Regression (LR)	*C*	0,1	Mengontrol kekuatan regularisasi
	penalty	12	Jenis regularisasi yang digunakan
	solver	lbfgs	Algoritma optimasi untuk menyelesaikan fungsi *log-loss*
K-NN	*n_neighbors*	11	Jumlah tetangga terdekat yang digunakan dalam klasifikasi
	weights	distance	Bobot diberikan lebih besar pada tetangga yang lebih dekat
Decision Tree (DT)	*max_depth*	3	Batas maksimum kedalaman pohon keputusan
	min_samples_split	2	Minimum jumlah sampel untuk membagi simpul
Random Forest (RF)	*n_estimators*	100	Jumlah pohon
	max_depth	none	Kedalaman pohon
	min_samples_split	5	Minimum jumlah sampel untuk membagi simpul pada setiap pohon
SVM	*C*	0,1	Mengatur kompromi antara margin maksimal dan kesalahan klasifikasi
	kernel	linear	Fungsi kernel yang digunakan untuk memetakan ke ruang dimensi lebih tinggi
	gamma	scale	Parameter kernel

Setelah diperoleh kombinasi *hyperparameter* terbaik dari proses *tuning*, setiap model kemudian diuji kembali menggunakan konfigurasi optimal tersebut. Pengujian dilakukan terhadap data uji yang telah disesuaikan dengan metode normalisasi masing-masing, guna menghasilkan prediksi yang merepresentasikan performa akhir dari model. Hasil prediksi inilah yang kemudian akan digunakan dalam tahap evaluasi selanjutnya.

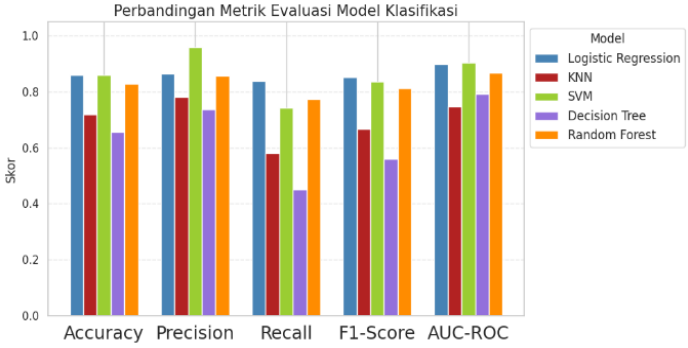
3.7 Hasil Evaluasi

Setelah model diuji ulang menggunakan parameter terbaik dari hasil tuning *hyperparameter*, diperoleh hasil evaluasi model serta analisis perbandingan performa antaralgoritma. Hasil evaluasi pada data uji berdasarkan lima metrik utama, yaitu akurasi, precision, *recall*, *F1-score*, dan AUC-ROC. Hasil evaluasi model machine learning bisa dilihat pada Tabel 5 sebagai berikut.

Tabel 5. Hasil Evaluasi Model

Model	*Accuracy*	*Precision*	*Recall*	*F1-Score*	AUC-ROC
Logistic Regression	0.859	0.867	0.839	0.852	0.900
K-Nearest Neighbors	0.719	0.783	0.581	0.667	0.749
Support Vector Machine	0.859	0.958	0.742	0.836	0.904
Decision Tree	0.656	0.737	0.452	0.560	0.793
Random Forest	0.828	0.857	0.774	0.814	0.868

Berdasarkan hasil evaluasi model pada Tabel 5, model *Support Vector Machine* dan *Logistic Regression* menunjukkan performa tertinggi dengan akurasi sebesar 85,9% serta nilai AUC-ROC di atas 0,90, yang menandakan kemampuan yang sangat baik dalam membedakan antara penderita dan non-penderita batu empedu. SVM memiliki nilai presisi tertinggi sebesar 95,83%, menunjukkan kemampuannya yang sangat baik dalam memprediksi kelas positif secara tepat. Sementara itu, *Logistic Regression* memiliki nilai *recall* tertinggi dibandingkan dengan model dengan akurasi serupa. Model *Random Forest* juga menunjukkan hasil yang kompetitif, dengan akurasi 82,8% dan keseimbangan yang baik antara *precision* dan *recall*. Sebaliknya, model *Decision Tree* dan K-NN menunjukkan performa yang lebih rendah, terutama pada metrik *recall*, yang menunjukkan banyaknya kasus positif yang tidak berhasil dikenali dengan baik oleh model. Hasil perbandingan metrik evaluasi model klasifikasi dapat divisualisasikan dalam bentuk grafik pada Gambar 5 sebagai berikut:



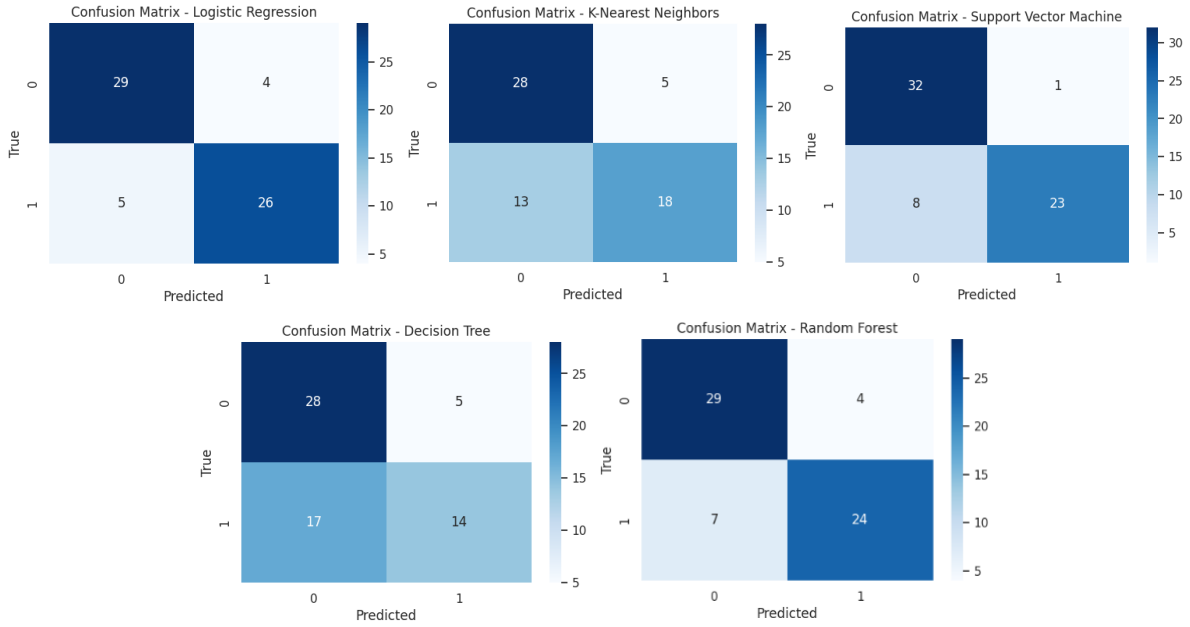
Gambar 5. Perbandingan Model *Machine Learning*

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

Dari hasil perbandingan model Machine Learning pada Gambar 5, dapat disimpulkan bahwa algoritma *Logistic Regression* dan SVM merupakan algoritma klasifikasi dengan nilai tertinggi dan menjadi algoritma klasifikasi yang baik untuk prediksi penyakit batu empedu berdasarkan data klinis dan biologis pasien.

3.8 Confusion Matrix

Model Confusion Matrix akan membentuk matriks yang terdiri dari true positive, false positive, true negative dan false negative. Berikut ini merupakan hasil dari confusion matrix dari kelima algoritma klasifikasi yang digunakan. Hasil perhitungan akurasi lima algoritma dapat dilihat pada grafik heatmap pada Gambar 6 berikut.



Gambar 6. Hasil perbandingan akurasi pada model Machine Learning

Pada Gambar 6 dapat dijelaskan sebagai berikut:

1. ***Confusion Matrix Logistic Regression***

Confusion matrix menunjukkan bahwa sebanyak 29 data penderita (label 0) berhasil diklasifikasikan dengan benar (True Positive), dan 26 data non-penderita (label 1) juga berhasil diklasifikasikan dengan benar (True Negative). Namun demikian, terdapat 4 data penderita yang salah diklasifikasikan sebagai non-penderita (False Positive), serta 5 data non-penderita yang salah diprediksi sebagai penderita (False Negative).

2. ***Confusion Matrix K-Nearest Neighbor***

Confusion matrix menunjukkan bahwa sebanyak 28 data penderita (label 0) berhasil diklasifikasikan dengan benar (True Positive), dan 18 data non-penderita (label 1) juga berhasil diklasifikasikan dengan benar (True Negative). Namun demikian, terdapat 5 data penderita yang salah diklasifikasikan sebagai non-penderita (False Positive), serta 13 data non-penderita yang salah diprediksi sebagai penderita (False Negative).

3. ***Confusion Matrix Support Vector Machine***

Confusion matrix menunjukkan bahwa sebanyak 32 data penderita (label 0) berhasil diklasifikasikan dengan benar (True Positive), dan 23 data non-penderita (label 1) juga berhasil diklasifikasikan dengan benar (True Negative). Namun demikian, terdapat 1 data penderita yang salah diklasifikasikan sebagai non-penderita (False Positive), serta 8 data non-penderita yang salah diprediksi sebagai penderita (False Negative).

4. ***Confusion Matrix Decision Tree***

Confusion matrix menunjukkan bahwa sebanyak 28 data penderita (label 0) berhasil diklasifikasikan dengan benar (*True Positive*), dan 14 data non-penderita (label 1) juga berhasil diklasifikasikan dengan benar (*True Negative*). Namun demikian, terdapat 5 data penderita yang salah diklasifikasikan sebagai non-penderita (*False Positive*), serta 17 data non-penderita yang salah diprediksi sebagai penderita (*False Negative*).

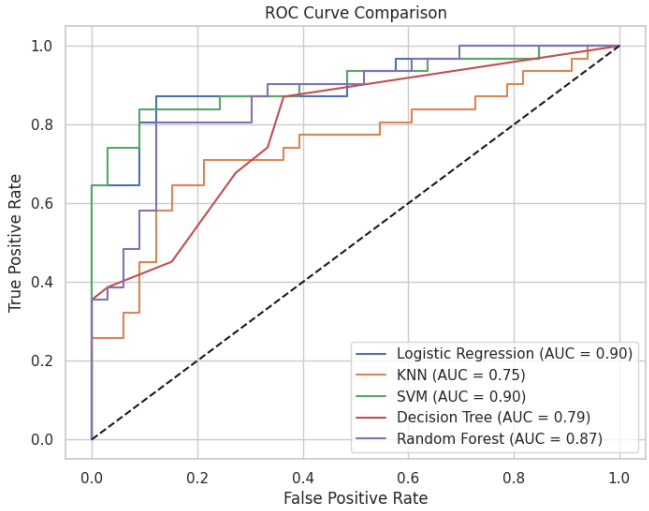
5. ***Confusion Matrix Random Forest***

Confusion matrix menunjukkan bahwa sebanyak 29 data penderita (label 0) berhasil diklasifikasikan dengan benar (True Positive), dan 24 data non-penderita (label 1) juga berhasil diklasifikasikan dengan benar (True Negative). Namun demikian, terdapat 4 data penderita yang salah diklasifikasikan sebagai non-penderita (False Positive), serta 7 data non-penderita yang salah diprediksi sebagai penderita (False Negative).

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

3.9 Kurva ROC dan Nilai AUC

Kurva ROC menggambarkan hubungan antara *True Positive Rate* (tingkat keberhasilan model dalam mengenali kelas positif) dengan *False Positive Rate* (tingkat kesalahan dalam mengklasifikasikan kelas negatif sebagai positif) pada berbagai ambang batas klasifikasi. Kurva ROC dapat dilihat pada Gambar 7 sebagai berikut.



Gambar 7. Kurva ROC

Pada Gambar 7 dapat dijelaskan perbandingan kurva ROC (*Receiver Operating Characteristic*) dari lima algoritma klasifikasi yang digunakan dalam penelitian ini. Model yang ideal ditandai dengan kurva yang melengkung tajam ke arah kiri atas, yang menunjukkan kemampuan model untuk membedakan kelas dengan baik, bahkan pada ambang yang bervariasi. Dari grafik tersebut terlihat bahwa model *Logistic Regression* dan *Support Vector Machine* (SVM) menunjukkan performa terbaik dengan nilai Area Under the Curve (AUC) sebesar 0.90. Nilai AUC ini mencerminkan tingkat akurasi keseluruhan yang tinggi dalam membedakan antara penderita dan non-penderita batu empedu. Random Forest menyusul dengan nilai AUC sebesar 0.87, yang juga menunjukkan performa yang cukup baik meskipun sedikit lebih rendah dibandingkan dengan dua model sebelumnya. Di sisi lain, Decision Tree menghasilkan AUC sebesar 0.79 dan *K-Nearest Neighbors* (KNN) sebesar 0.75, yang menandakan bahwa kedua model ini kurang stabil dan cenderung memiliki kemampuan pemisahan kelas yang lebih lemah dibandingkan dengan model-model lainnya. Secara umum, grafik ini memperkuat temuan sebelumnya bahwa Logistic Regression dan SVM merupakan model yang paling andal dalam prediksi penyakit batu empedu.

3.10 Pembahasan

Pada bagian ini akan disajikan evaluasi perbandingan hasil temuan penelitian terhadap penelitian sebelumnya. Hasil evaluasi bisa dilihat pada Tabel 6 sebagai berikut.

Tabel 6. Temuan Hasil Penelitian Terhadap Penelitian Sebelumnya

Penelitian	*Model*	*Accuracy*	*Precision*	*Recall*	*F1-Score (%)*	AUC-ROC (%)
[8]	*Gradient Boosted Tree*	82.29	89.63	72.58	-	-
[1]	Teorema Bayes	69.02	-	-	-	
[9]	Random Forest dan Support Vector Machine	98.00	96.00	100	98.00	-
[10]	K-Nearest Neighbor	92.00	90.00	92.00	-	
Hasil Penelitian saat ini	Logistic Regression	85.90	86.70	83.90	85.20	90.00
	K-Nearest Neighbors	71.90	78.30	58.10	66.70	74.90
	Support Vector Machine	85.90	95.80	74.20	83.60	90.40
	Decision Tree	65.60	73.70	45.20	56.00	79.30
	Random Forest	82.80	85.70	77.40	81.40	86.80

Berdasarkan Tabel 7, penelitian ini menunjukkan bahwa algoritma Logistic Regression dan Support Vector Machine (SVM) menghasilkan performa klasifikasi terbaik dengan nilai accuracy sebesar 85,90%. Meskipun memiliki nilai akurasi yang sama, SVM menunjukkan keunggulan pada nilai precision sebesar 95,80% dan AUC-ROC sebesar 90,40%, yang mengindikasikan kemampuan model dalam membedakan kelas

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

positif dan negatif secara lebih baik. Di sisi lain, Logistic Regression memperoleh nilai recall sebesar 83,90%, lebih tinggi dibandingkan SVM, sehingga memiliki kemampuan yang lebih baik dalam mengidentifikasi kasus penderita batu empedu. Temuan ini menunjukkan bahwa kedua algoritma memiliki karakteristik yang berbeda, di mana SVM lebih unggul dalam meminimalkan *false positive*, sedangkan Logistic Regression lebih efektif dalam mendeteksi kasus positif.

Jika dibandingkan dengan penelitian yang menggunakan Gradient Boosted Tree, penelitian ini menghasilkan akurasi yang lebih tinggi, yaitu 85,90% dibandingkan dengan 82,29%. Selain itu, nilai recall Logistic Regression (83,90%) juga lebih tinggi dibandingkan dengan penelitian tersebut (72,58%), sehingga menunjukkan peningkatan kemampuan model dalam mendeteksi penderita batu empedu. Dibandingkan dengan penelitian yang menerapkan K-Nearest Neighbor, penelitian tersebut memperoleh akurasi sebesar 92,00%, lebih tinggi dibandingkan dengan hasil K-Nearest Neighbor pada penelitian ini yang mencapai 71,90%. Perbedaan performa tersebut diduga dipengaruhi oleh karakteristik dataset, jumlah fitur, teknik prapemrosesan data, proses optimasi *hyperparameter*, serta pembagian data latih dan data uji yang berbeda pada masing-masing penelitian.

Sementara itu, penelitian yang menggunakan Random Forest dan Support Vector Machine melaporkan performa yang sangat tinggi dengan accuracy sebesar 98,00%, precision 96,00%, recall 100%, dan F1-score 98,00%. Meskipun hasil tersebut lebih tinggi dibandingkan dengan penelitian ini, perbandingan secara langsung tidak dapat dilakukan karena adanya perbedaan dataset, karakteristik data, teknik prapemrosesan, serta konfigurasi model yang digunakan. Faktor-faktor tersebut diketahui berpengaruh terhadap kemampuan generalisasi model machine learning.

Secara keseluruhan, penelitian ini memberikan kontribusi melalui evaluasi komprehensif terhadap lima algoritma *machine learning* pada dataset penyakit batu empedu dengan menerapkan Winsorized Transformation, RobustScaler, dan optimasi *hyperparameter* menggunakan GridSearchCV. Selain itu, evaluasi dilakukan menggunakan lima metrik, yaitu accuracy, precision, recall, F1-score, dan ROC AUC, sehingga memberikan gambaran performa model yang lebih menyeluruh dibandingkan dengan beberapa penelitian terdahulu yang hanya menggunakan sebagian metrik evaluasi.

4. KESIMPULAN

Penelitian ini bertujuan untuk membandingkan kinerja lima algoritma klasifikasi, yaitu Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), dan K-Nearest Neighbors (K-NN) dalam memprediksi penderita penyakit batu empedu berdasarkan data klinis dan biologis pasien. Dataset yang digunakan memiliki kualitas yang baik dan seimbang antara penderita dan non-penderita, serta terdiri dari 38 fitur klinis numerik. Untuk menangani outlier, digunakan pendekatan normalisasi berbasis Winsorized Transformation, RobustScaler, serta kombinasi keduanya sesuai karakteristik masing-masing algoritma. Proses modeling dilakukan dengan tuning hyperparameter melalui GridSearchCV dan evaluasi performa menggunakan metrik akurasi, presisi, recall, F1-score, serta ROC AUC. Hasil penelitian menunjukkan bahwa model Support Vector Machine dan Logistic Regression memiliki performa terbaik dengan akurasi sebesar 85,9% dan nilai AUC di atas 0,90. Random Forest juga memberikan hasil yang kompetitif dengan akurasi 82,8%. Sebaliknya, Decision Tree dan K-NN menunjukkan performa yang lebih rendah, terutama pada recall. Berdasarkan hasil tersebut, dapat disimpulkan bahwa Logistic Regression dan SVM merupakan algoritma yang paling andal dan akurat untuk prediksi penyakit batu empedu pada data klinis non-imaging. Kontribusi utama penelitian ini adalah memberikan evaluasi komprehensif terhadap lima algoritma machine learning dengan menerapkan strategi penanganan outlier yang disesuaikan dengan karakteristik algoritma, sehingga dapat menjadi acuan dalam pemilihan model klasifikasi untuk prediksi penyakit batu empedu berbasis data klinis non-imaging. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena menggunakan satu sumber dataset publik (UCI Machine Learning Repository), sehingga karakteristik data belum sepenuhnya merepresentasikan populasi pasien yang lebih luas. Selain itu, penelitian ini hanya mengevaluasi algoritma machine learning konvensional tanpa membandingkannya dengan metode ensemble yang lebih kompleks atau model deep learning. Oleh karena itu, penelitian selanjutnya dapat memanfaatkan dataset dari berbagai institusi kesehatan, melakukan validasi eksternal, serta mengeksplorasi algoritma yang lebih mutakhir agar model yang dihasilkan memiliki kemampuan generalisasi yang lebih baik dan siap diimplementasikan pada lingkungan klinis.

REFERENCES

[1] M. Riandini, S. Wardani, and M. Handayani, "Penerapan Metode Teorema Bayes pada Sistem Pakar dalam Mendiagnosa Tingkat Kepastian Penyakit Batu Empedu," *J. Teknol. Sist. Inf. dan Sist. Komput. TGD*, vol. 8, no. 2, pp. 132–139, 2025, doi: 10.53513/jsk.v8i2.11465.

[2] R. R. Hidayah and T. Y. Pramana, "Peran Ursodeoxycholic Acid (Udca) Dalam Pengelolaan Penyakit Batu Empedu," *Prepotif J. Kesehat. Masy.*, vol. 9, no. 1, pp. 2357–2368, 2025, doi: 10.31004/prepotif.v9i1.43148.

[3] A. H. Andini, M. Romdhoni, and F. Oktavrisa, "Karakteristik Pasien Batu Empedu Yang di Rawat di RSUD Waled Periode 2019-2022," *Syntax Lit. ; J. Ilm. Indones.*, vol. 7, no. 9, pp. 15291–15303, 2023, doi: 10.36418/syntax-literate.v7i9.14253.

Bulletin of Information System Research (BIOS)
Volume 4, No 2, April 2026, Page: 90−101
ISSN 2963-2455 (media online)
DOI: 10.62866/bios.v4i2.281
https://journal.grahamitra.id/index.php/bios

[4] B. Fitriyasti, Khomaini, R. Aulia, and S. Ferilda, "Karakteristik Pasien Kolelitiasis Di Rumah Sakit Siti Rahmah Padang Pada Tahun 2022," *J. Kesehat. Saintika Meditory*, vol. 8, no. 1, pp. 405–414, 2025, doi: 10.30633/jsm.v8i1.3231.
[5] I. Akbar, F. Supriadi, and D. Indra Junaedi, "Pemanfaatan Machine Learning Di Bidang Kesehatan," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 1, pp. 1744–1749, 2025, doi: 10.36040/jati.v9i1.12663.
[6] R. M. Siregar, B. Satria, S. Fadilah, L. Mayola, and S. Safira, "SMOTE-Based Comparative Analysis of Machine Learning Models for Stroke Risk Prediction Using Imbalanced Healthcare Data," *Ilk. J. Ilm.*, vol. 18, no. 1, pp. 180–194, 2026, doi: 10.33096/ilkom.v18i1.3161.180-194.
[7] Fadlina, "Data Mining Untuk Prediksi Penyakit Batu Empedu Dengan Algoritma C45 Aplikasi Rapid Miner," *J. Gemilang Inform.*, vol. 1, no. 2, 2024, doi: 10.58369/git.v1i2.126.
[8] F. G. Amali, H. Firmansyah, and W. Asriani, "Evaluasi Kinerja Model Gradient Boosted Trees Untuk Prediksi Status Komorbiditas Pada Pasien Batu Empedu," *J. Sist. Inf. dan Sains Teknol.*, vol. 8, no. 1, pp. 21–37, 2026, doi: 10.31326/sistek.v8i1.2641.
[9] A. Ghozali, H. Pratiwi, and S. S. Handajani, "Implementasi Data Mining Menggunakan Metode Random Forest Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes," *Delta J. Ilm. Pendidik. Mat.*, vol. 11, no. 2, pp. 147–162, 2023, doi: 10.31941/delta.v11i2.2686.
[10] M. F. Akbarollah, W. Wiyanto, D. Ardiatma, and A. T. Zy, "Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Penyakit Jantung," *J. Comput. Syst. Informatics*, vol. 4, no. 4, pp. 850–860, 2023, doi: 10.47065/josyc.v4i4.4071.
[11] H. Hidayat, M. F. Wicaksono, Y. J. Wicaksono, S. M. Wibawa, and E. N. P. Septia, "Prediksi Penyakit Liver Dengan Model Pebandingan Logistic Regression, Support Vector Machine, Dan Random Forest," *Komputika J. Sist. Komput.*, vol. 15, no. 1, pp. 83–94, 2026, doi: 10.34010/komputika.v15i1.19799.
[12] A. D. Achmad, "Klasifikasi Breast Cancer Menggunakan Metode Logistic Regression," *JTRISTE*, vol. 9, no. 1, 2022.
[13] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naïve Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, 2021, doi: 10.32493/informatika.v5i4.7622.
[14] B. Satria, N. Afrianto, L. Ningsih, P. Sakinah, A. Sidauruk, and L. Mayola, "Comparative Analysis of Weighted-KNN, Random Forest, and Support Vector Machine Models for Beef and Pork Image Classification Using Machine Learning," *Int. J. Informatics Vis.*, vol. 9, no. 4, pp. 1677–1687, 2025, doi: 10.62527/joiv.9.4.3736.
[15] A. Primajaya and B. N. Sari, "Random Forest Algorithm for Prediction of Precipitation," *Indones. J. Artif. Intell. Data Min.*, vol. 1, no. 1, 2018, doi: 10.24014/ijaidm.v1i1.4903.
[16] M. Premalatha and C. Vijayalakshmi, "SVM approach for non-parametric method in classification and regression learning process on feature selection with ε-insensitive region," *Malaya J. Mat.*, vol. S, no. 1, pp. 276–279, 2019, doi: 10.26637/MJM0S01/0051.
[17] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, 1995, doi: 10.1007/bf00994018.
[18] D. Cahyanti, A. Rahmayani, and S. A. Husniar, "Analisis performa metode Knn pada Dataset pasien pengidap Kanker Payudara," *Indones. J. Data Sci.*, vol. 1, no. 2, 2020, doi: 10.33096/ijodas.v1i2.13.
[19] Erlin, Yulvia Nora Marlim, Junadhi, Laili Suryati, and Nova Agustina, "Deteksi Dini Penyakit Diabetes Menggunakan Machine Learning dengan Algoritma Logistic Regression," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 88–96, 2022, doi: 10.22146/jnteti.v11i2.3586.
[20] M. Z. Nurrokhman, "Perbandingan Algoritma Support Vector Machine dan Neural Network untuk Klasifikasi Penyakit Hati Ma'mur," *Indones. J. Comput. Sci.*, vol. 12, no. 4, pp. 2096–2106, 2023, doi: 10.33022/ijcs.v12i4.3274.