

# Optimasi *Hyperparameter Ensemble Learning* untuk Prediksi Penyakit *Liver* Berdasarkan Data Pasien

Agung Baitul Hikmah

Universitas Bina Sarana Informatika, Jl. Kramat Raya No.98, Kwitang, Kec. Senen, DKI Jakarta  
[agung.abl@bsi.ac.id](mailto:agung.abl@bsi.ac.id)

Diterima : 25 April 2025

Disetujui : 26 Mei 2025

**Abstract**—Penyakit *liver* merupakan salah satu masalah kesehatan serius yang memerlukan deteksi dini guna meningkatkan peluang pengobatan yang efektif. Penelitian ini bertujuan untuk mengembangkan model prediksi penyakit *liver* berdasarkan data pasien dengan menggunakan teknik *ensemble learning*, yaitu Random Forest, XGBoost, AdaBoost, dan Extra Trees Classifier. Dataset yang digunakan mencakup berbagai parameter medis pasien yang berkontribusi terhadap diagnosis penyakit *liver*. Evaluasi model dilakukan menggunakan metrik akurasi, AUC, recall, precision, dan F1-score. Hasil penelitian menunjukkan bahwa *Extra Trees Classifier* memiliki performa terbaik, dengan akurasi sebesar 73.61%, AUC 76.82%, recall 89.82%, precision 77.22%, dan F1-score 82.93%. Model ini memiliki kemampuan yang sangat baik dalam mengidentifikasi *pasien liver* (*recall* tinggi), serta keseimbangan antara *precision* dan *recall* yang optimal. Selain itu, analisis menggunakan *confusion matrix* menunjukkan bahwa model ini mampu memprediksi sebagian besar kasus positif dengan benar, meskipun masih terdapat beberapa kesalahan klasifikasi pada pasien *non-liver*. Kurva AUC-ROC mengkonfirmasi bahwa model ini cukup andal dalam membedakan antara pasien *liver* dan *non-liver*, dengan nilai *macro-average* AUC sebesar 0.81 dan *micro-average* AUC 0.85. Berdasarkan hasil ini, *Extra Trees Classifier* direkomendasikan sebagai model terbaik untuk sistem pendukung keputusan dalam diagnosis *penyakit liver*. Untuk meningkatkan performa model di masa depan, diperlukan penggunaan *dataset* yang lebih besar, optimasi *hyperparameter* lebih lanjut, serta eksplorasi teknik *balancing* data untuk meningkatkan akurasi klasifikasi.

**Keywords** — Penyakit *Liver*, *Ensemble Learning*, *Extra Tree Classification*, *Machine Learning*, Prediksi.

## I. PENDAHULUAN

Penyakit *liver* merupakan salah satu masalah kesehatan global yang memiliki dampak signifikan terhadap kualitas hidup masyarakat [1], [2]. Penyakit *liver*, yang meliputi berbagai jenis gangguan pada organ hati seperti hepatitis, sirosis, dan kanker hati, merupakan salah satu penyebab utama kematian di seluruh dunia [3]. Organ hati memiliki peran vital dalam tubuh manusia, seperti dalam proses detoksifikasi, metabolisme, dan produksi protein penting [1], [4]. Gangguan pada fungsi hati dapat berakibat fatal jika tidak segera didiagnosis dan ditangani dengan baik [5]. Di Indonesia, kasus penyakit *liver* semakin meningkat, terutama pada individu yang terpapar faktor risiko seperti infeksi virus

hepatitis, konsumsi alkohol berlebihan, serta pola makan yang tidak sehat [6]. Oleh karena itu, deteksi dini penyakit *liver* sangat penting untuk mencegah komplikasi yang lebih serius [2].

Proses diagnosis penyakit *liver* seringkali melibatkan serangkaian pemeriksaan medis yang kompleks dan memerlukan keahlian khusus dari tenaga medis, seperti tes darah, biopsi hati, dan pencitraan medis [7]. Selain memakan waktu, prosedur ini juga dapat meningkatkan risiko keterlambatan diagnosis dan pengobatan, karena hasil diagnosis sangat bergantung pada pengalaman dan keahlian tenaga medis [8], [9]. Oleh karena itu, diperlukan sistem yang lebih objektif dan efisien untuk mendukung proses diagnosis penyakit *liver* [10]. Salah satu

pendekatan yang menjanjikan adalah penggunaan teknologi kecerdasan buatan, khususnya data mining dan machine learning, yang dapat meningkatkan akurasi dan efisiensi dalam diagnosis penyakit *liver* [11].

Algoritma *machine learning* dapat membantu menganalisis data pasien dan mendeteksi pola yang mungkin tidak dapat ditemukan oleh tenaga medis secara manual [6]. Di antara berbagai algoritma yang ada, *ensemble learning* telah terbukti memberikan hasil yang lebih akurat dalam berbagai masalah prediksi. Teknik *ensemble* menggabungkan hasil dari beberapa model pembelajaran untuk menghasilkan prediksi yang lebih baik dibandingkan dengan model tunggal [12], [13]. Metode ini mengurangi variansi dan bias, serta dapat mengoptimalkan hasil prediksi dalam kondisi yang kompleks, seperti pada data medis yang sering kali memiliki ketidakpastian dan ketidaksempurnaan [14].

## II. PENELITIAN TERKAIT

Penelitian tentang algoritma *ensemble*, seperti *Random Forest*, *Gradient Boosting*, dan *AdaBoost*, telah digunakan dalam berbagai aplikasi medis, termasuk prediksi penyakit jantung, kanker, dan diabetes [13]. Keunggulan dari algoritma ini adalah kemampuannya untuk bekerja dengan berbagai jenis data dan menangani kompleksitas yang ada [12]. Pada penelitian ini, penulis akan menerapkan teknik *ensemble learning* untuk memprediksi penyakit *liver* berdasarkan data pasien yang meliputi berbagai fitur seperti usia, jenis kelamin, riwayat penyakit, serta hasil tes medis. Diharapkan dengan menggunakan pendekatan ini, dapat diperoleh model yang akurat dalam memprediksi kemungkinan seorang pasien menderita penyakit *liver*.

Penelitian yang mengkaji penggunaan algoritma supervised learning dalam klasifikasi penyakit *liver* membahas berbagai metode klasifikasi, termasuk Naïve Bayes, K-Nearest Neighbors (KNN), Decision Tree, dan Neural Network, yang diterapkan pada dataset *Indian Liver Patient Dataset* (ILPD) dari UCI Machine Learning Repository [8]. Hasil penelitian

menunjukkan bahwa algoritma *Decision Tree* mencapai akurasi terbaik.

Selanjutnya penelitian oleh [11] studi kasus *liver* ini menggunakan model KNN untuk memprediksi apakah seorang pasien menderita penyakit *liver* atau tidak. Penelitian ini mencatat bahwa KNN dapat terpengaruh oleh dimensi data yang tinggi dan noise dalam data, yang dapat mempengaruhi kinerja model. Oleh karena itu, penting untuk melakukan pemilihan parameter dengan hati-hati, seperti nilai  $k$ , untuk meningkatkan akurasi.

Perbaikan dari penelitian sebelumnya dilakukan oleh [4] dengan menerapkan salah satu algoritma *ensemble learning* yaitu random forest. Hasil penelitian menunjukkan bahwa Random Forest memberikan hasil yang cukup baik dalam memprediksi penyakit *liver* dengan akurasi yang tinggi. Penelitian ini juga menggarisbawahi keunggulan Random Forest dalam menangani data yang kompleks dan tidak seimbang, seperti dataset penyakit *liver* yang sering kali memiliki banyak data yang tidak seimbang.

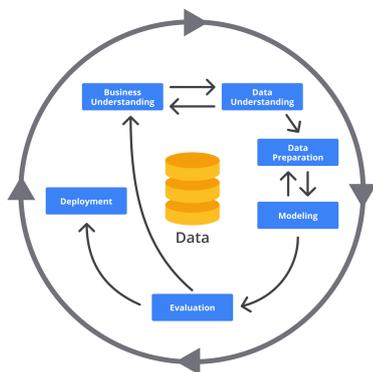
Meskipun banyak penelitian yang telah menggunakan *machine learning* untuk prediksi penyakit *liver*, sebagian besar fokus pada penggunaan algoritma tunggal tanpa eksplorasi yang mendalam terhadap algoritma *ensemble*. Berdasarkan tinjauan literatur terkini, penggunaan *ensemble methods* dalam prediksi berbagai penyakit menunjukkan potensi untuk meningkatkan akurasi model dengan menggabungkan keunggulan beberapa algoritma. Sebagai contoh penelitian oleh [15], menunjukkan bahwa algoritma *ensemble* dapat mengurangi variansi dan bias dalam model prediksi, yang memungkinkan akurasi yang lebih tinggi dalam diagnosis penyakit. Namun, penelitian yang secara spesifik mengoptimalkan penggunaan algoritma *ensemble* untuk prediksi penyakit *liver* masih relative [16].

Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi beberapa algoritma *ensemble* dalam prediksi penyakit *liver*, serta membandingkan kinerja masing-masing model dengan pendekatan yang ada sebelumnya. Pemilihan algoritma *ensemble* didasarkan pada kemampuannya dalam mengurangi bias dan

varians, serta meningkatkan generalisasi model yang menjadi dasar penting dalam prediksi penyakit medis [15]. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem prediksi penyakit *liver* yang lebih cepat dan akurat serta memberikan sumber referensi bagi penerapan algoritma *ensemble* dalam dunia medis, sehingga dapat digunakan sebagai alat bantu bagi tenaga medis dalam proses diagnosis.

### III. METODE PENELITIAN

Dalam penelitian ini, digunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*) sebagai pendekatan yang umum dalam menyelesaikan masalah bisnis dan penelitian [17]. Metodologi ini terdiri dari enam tahapan yang saling terhubung, yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment* [18]. Setiap tahapan dalam CRISP-DM memiliki peran penting yang mendukung proses analisis data, mulai dari pemahaman terhadap tujuan bisnis hingga penerapan model yang dihasilkan [19]. Alur CRISP-DM digambarkan pada Gambar 1.



Gambar 1. Metode CRISP-DM

Pada gambar 1 menunjukkan proses yang sistematis ini memungkinkan penelitian untuk menghasilkan solusi yang akurat dan dapat diandalkan [20]. Dalam penelitian ini, pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*) digunakan sebagai metodologi untuk mengembangkan model prediksi penyakit *liver* menggunakan algoritma *ensemble learning*.

#### A. Business Understanding

Tahap ini bertujuan untuk memahami permasalahan yang akan diselesaikan dan menentukan tujuan penelitian. Penyakit *liver* merupakan masalah kesehatan serius yang membutuhkan metode deteksi dini yang lebih akurat. Oleh karena itu, penelitian ini bertujuan untuk menerapkan metode *ensemble learning* guna meningkatkan akurasi prediksi penyakit *liver* berdasarkan data pasien. Model yang dikembangkan diharapkan dapat membantu tenaga medis dalam mengambil keputusan yang lebih cepat dan akurat [12].

#### B. Data Understanding

Setelah memahami permasalahan, langkah selanjutnya adalah mengumpulkan dan memahami data pasien yang akan digunakan. Dataset yang digunakan mencakup fitur-fitur seperti usia, kadar enzim hati (SGPT, SGOT), kadar albumin, total protein, serta indikator lain yang relevan dengan kondisi kesehatan *liver*. Pada tahap ini, dilakukan eksplorasi data untuk memahami distribusi nilai, mengidentifikasi data yang hilang (*missing values*), serta melihat korelasi antara variabel yang ada dalam dataset.

#### C. Data Preprocessing

Tahap ini berfokus pada persiapan data agar dapat digunakan dalam pemodelan. Data yang telah dikumpulkan akan dibersihkan dengan menghapus atau menangani nilai yang hilang dan menghilangkan outlier yang dapat mempengaruhi hasil prediksi. Selanjutnya, proses normalisasi atau standardisasi dilakukan agar skala data menjadi seragam. Data juga akan dibagi menjadi *training set* dan *testing set* untuk memastikan model yang dibangun dapat dievaluasi dengan baik.

#### D. Modeling

Pada tahap ini, model *ensemble learning* diterapkan untuk membangun sistem prediksi penyakit *liver*. Algoritma yang digunakan meliputi Random Forest, XGBoost, AdaBoost dan Extra tree yang dikenal memiliki performa tinggi dalam klasifikasi. Model akan diuji dengan berbagai parameter untuk menentukan

kombinasi terbaik yang memberikan hasil paling akurat.

1) *Random Forest*

Random Forest bekerja berdasarkan konsep *ensemble learning*, yang menggabungkan *multiple decision trees* untuk menghasilkan prediksi yang lebih akurat dan stabil [21]. Secara matematis, proses dalam Random Forest untuk menentukan prediksi akhir  $\hat{y}$  diuraikan pada persamaan 1.

$$\hat{y} = mode (h1(X), h2 (X), \dots, hT(X)) \quad (1)$$

$\hat{y}$  adalah Prediksi Akhir  
 $h1(X)$  adalah pohon ke-i

2) *XGBoost*

XGBoost (*Extreme Gradient Boosting*) adalah algoritma yang digunakan dalam tugas klasifikasi untuk membangun sekumpulan *decision trees* yang diperbarui secara iteratif untuk meminimalkan kesalahan prediksi [22]. Algoritma ini sangat populer karena kecepatan, efisiensi memori, dan kemampuannya menangani dataset besar dengan fitur yang kompleks [23]. Persamaan XGBoost Klasifikasi.

$$L(y, \hat{y}) = \sum_{i=1}^n [(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (2)$$

$y$  = label asli (*ground truth*), bernilai 0 atau 1  
 $\hat{y}$  = probabilitas prediksi (nilai antara 0 dan 1)

3) *Adaboost*

AdaBoost (*Adaptive Boosting*) adalah algoritma *ensemble learning* yang menggabungkan beberapa *weak classifiers* (biasanya *decision trees* dengan  $depth = 1$ , disebut *stumps*) untuk membentuk model yang lebih kuat. Setiap iterasi meningkatkan bobot pada sampel yang sulit diklasifikasikan [24], sehingga model lebih fokus pada kesalahan sebelumnya. Persamaan Adaboost Klasifikasi.

$$H(x) = sign (\sum_{t=1}^T at . ht(x)) \quad (3)$$

$H(x)$  adalah label kelas akhir

4) *Extra Tree*

*Extra Trees (Extremely Randomized Trees)* adalah varian dari Random Forest yang membangun sekumpulan *decision trees* untuk tugas klasifikasi dan regresi [21]. Berbeda dengan Random Forest, Extra Trees memilih fitur dan nilai split secara acak, tanpa mencari pemisahan terbaik berdasarkan informasi gain atau Gini impurity [25].

$$H(x) = mode (h1(x), h2 (x), \dots, hT(x)) \quad (4)$$

$H(x)$  = prediksi akhir dari model *ensemble*.  
 $(h1(x), h2 (x), \dots, hT(x))$ , adalah prediksi dari masing-masing  $T$  model individu dalam *ensemble*.

E. *Evaluation*

Setelah model dibuat, tahap evaluasi dilakukan untuk menilai kinerjanya. Metrik yang digunakan mencakup akurasi, precision, recall, F1-score, dan AUC-ROC, yang dapat mengukur sejauh mana model mampu memprediksi penyakit *liver* dengan benar. Hasil evaluasi ini juga akan dibandingkan antara berbagai algoritma *ensemble* yang digunakan untuk menentukan model terbaik yang akan diterapkan.

F. *Deployment*

Jika model yang dikembangkan menunjukkan performa yang baik, tahap implementasi dapat dilakukan. Model dapat diterapkan dalam sistem pendukung keputusan medis untuk membantu dokter dalam menganalisis risiko penyakit *liver* berdasarkan data pasien. Selain itu, interpretasi hasil model juga disiapkan agar dapat digunakan secara praktis oleh tenaga medis dalam proses diagnosis.

IV. HASIL DAN PEMBAHASAN

Pada bagian ini, hasil eksperimen terkait optimasi hyperparameter pada berbagai algoritma *ensemble learning* untuk prediksi penyakit *liver* berdasarkan data pasien akan dibahas secara mendetail. Model yang diuji meliputi *Random Forest*, *XGBoost*, *AdaBoost*, dan *Extra Trees Classifier*, dengan evaluasi menggunakan metrik akurasi, AUC, recall, precision, dan F1-score.

Optimasi hyperparameter dilakukan untuk meningkatkan performa masing-masing model dalam membedakan pasien dengan dan tanpa penyakit *liver*. Analisis hasil dilakukan dengan membandingkan kinerja model sebelum dan sesudah optimasi, serta meninjau distribusi klasifikasi menggunakan confusion matrix dan kurva AUC-ROC. Selain itu, pembahasan akan mencakup pengaruh masing-masing fitur dalam proses prediksi serta strategi yang dapat diterapkan untuk lebih meningkatkan performa model.

#### A. Dataset Penelitian

Pada penelitian ini dataset yang digunakan bersumber dari data sekunder dari *Dataset Indian Liver Patient Dataset (ILPD)* berisi 583 data pasien dengan 11 fitur, yang digunakan untuk memprediksi apakah seseorang menderita penyakit *liver* atau tidak. Dataset ini terdiri dari berbagai atribut yang mencerminkan kondisi kesehatan pasien, termasuk usia, jenis kelamin, kadar bilirubin, protein, albumin, serta enzim hati seperti SGPT, SGOT, dan alkaline phosphatase (ALP) [26].

Berikut adalah gambaran umum dari data pasien yang diteliti, tabel 1 mendeskripsikan dari atribut-atribut yang digunakan pada penelitian ini.

Tabel 1. Gambaran umum dataset

Kolom	Deskripsi
age	Usia pasien (dalam tahun).
gender	Jenis kelamin pasien (Male/Female).
tot_bilirubin	Total kadar bilirubin dalam darah.
direct_bilirubin	Kadar bilirubin langsung dalam darah.
tot_proteins	Jumlah total protein dalam darah.
albumin	Konsentrasi albumin dalam darah.
ag_ratio	Rasio albumin terhadap globulin dalam darah.
sgpt	<i>Serum Glutamate Pyruvate Transaminase</i> (SGPT), enzim hati.
sgot	<i>Serum Glutamic-Oxaloacetic Transaminase</i> (SGOT), enzim hati.
alkphos	Kadar alkaline phosphatase (ALP), enzim hati.
is_patient	Label target: 1 untuk pasien <i>liver</i> , 2 untuk non-pasien.

Sample dataset penelitian dari beberapa pasien yang terindikasi *liver* atau tidak pada Gambar 2.

	age	gender	tot_bilirubin	direct_bilirubin	tot_proteins	albumin	ag_ratio	sgpt	sgot	alkphos	is_patient
0	65	Female	0.7	0.1	187	16	18	6.8	3.3	0.90	1
1	62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	1
2	62	Male	7.3	4.1	490	60	68	7.0	3.3	0.89	1
3	58	Male	1.0	0.4	182	14	20	6.8	3.4	1.00	1
4	72	Male	3.9	2.0	195	27	59	7.3	2.4	0.40	1
...	...	...	...	...	...	...	...	...	...	...	...
578	60	Male	0.5	0.1	500	20	34	5.9	1.6	0.37	2
579	40	Male	0.6	0.1	98	35	31	6.0	3.2	1.10	1
580	52	Male	0.8	0.2	245	48	49	6.4	3.2	1.00	1
581	31	Male	1.3	0.5	184	29	32	6.8	3.4	1.00	1
582	38	Male	1.0	0.3	216	21	24	7.3	4.4	1.50	2

Gambar 2. Sample dataset

Pada Gambar 2 menampilkan data pasien yang digunakan untuk analisis penyakit *liver*. Setiap baris mewakili satu individu dengan berbagai atribut kesehatan, termasuk usia, jenis kelamin, serta hasil uji laboratorium yang berkaitan dengan fungsi hati. Data ini mencakup parameter biokimia seperti kadar bilirubin, protein total, albumin, dan rasio albumin-globulin, yang sering digunakan untuk mengevaluasi kondisi hati seseorang. Selain itu, tabel juga mencantumkan enzim hati seperti SGPT, SGOT, dan alkaline phosphatase, yang merupakan indikator penting dalam mendeteksi gangguan hati. Kolom terakhir menunjukkan status pasien, di mana nilai 1 mengindikasikan apakah seseorang didiagnosis dengan penyakit *liver* atau nilai 2 tidak.

#### B. Preprocessing Dataset

Langkah selanjutnya adalah melakukan *preprocessing* data untuk memastikan bahwa dataset siap digunakan dalam analisis atau pemodelan machine learning. Tahapan *preprocessing* ini mencakup beberapa langkah penting, antara lain.

- Menangani Nilai yang Hilang (*Missing Values*). Dari hasil eksplorasi awal, ditemukan bahwa kolom *alkphos* (*alkaline phosphatase*) memiliki 4 nilai kosong. Nilai ini perlu diimputasi menggunakan median atau mean agar tidak mengganggu proses analisis data.
- Mengonversi Data Kategori ke Format Numerik. Kolom *gender* masih berbentuk teks (*Male/Female*) sehingga perlu dikonversi menjadi angka, misalnya 1 untuk Male dan 0 untuk Female agar dapat digunakan dalam model *machine learning*.

- Mengubah Label Target menjadi Biner. Kolom *is\_patient* memiliki nilai 1 dan 2, di mana 1 menunjukkan pasien *liver* dan 2 menunjukkan non-pasien. Agar lebih mudah digunakan dalam klasifikasi, label ini perlu diubah menjadi 0 (non-pasien) dan 1 (pasien *liver*).
- Menangani Outlier dan Distribusi Data, Beberapa fitur seperti total bilirubin, albumin, dan *ag\_ratio* memiliki nilai ekstrem yang jauh dari rata-rata. Outlier ini dapat ditangani dengan transformasi logaritmik atau normalisasi agar tidak mendominasi model.
- Pembagian Dataset, Dataset akan dibagi menjadi training set (80%) dan testing set (20%) agar model dapat diuji dengan data yang belum pernah dilihat sebelumnya [27].

Hasil dari tahapan *preprocessing* ditampilkan pada Gambar 3.

	age	gender	tot_bilirubin	direct_bilirubin	tot_proteins	albumin	ag_ratio	sgpt	sgot	alkphos	is_patient
0	65	0	0.7	0.1	187	16	18	6.8	3.3	0.90	1
1	62	1	10.9	5.5	699	64	100	7.5	3.2	0.74	1
2	62	1	7.3	4.1	490	60	68	7.0	3.3	0.89	1
3	58	1	1.0	0.4	182	14	20	6.8	3.4	1.00	1
4	72	1	3.9	2.0	195	27	59	7.3	2.4	0.40	1
...	...	...	...	...	...	...	...	...	...	...	...
578	60	1	0.5	0.1	500	20	34	5.9	1.6	0.37	0
579	40	1	0.6	0.1	98	35	31	6.0	3.2	1.10	1
580	52	1	0.8	0.2	245	48	49	6.4	3.2	1.00	1
581	31	1	1.3	0.5	184	29	32	6.8	3.4	1.00	1
582	38	1	1.0	0.3	216	21	24	7.3	4.4	1.50	0

583 rows x 11 columns

Gambar 3. Hasil Preprocessing Data

Pada Gambar 3 Dataset pasien telah diproses agar siap digunakan dalam model *machine learning*. Gender telah dikonversi ke numerik (1 untuk *Male*, 0 untuk *Female*), dan *is\_patient* telah diubah menjadi biner (1 untuk pasien *liver*, 0 untuk non-pasien). Nilai yang hilang telah diisi untuk memastikan kelengkapan data, sementara fitur seperti bilirubin, protein, enzim hati, dan *alkphos* telah dinormalisasi agar memiliki skala yang seragam. Outlier juga telah ditangani untuk menghindari pengaruh negatif pada model. Dengan perbaikan ini, dataset siap digunakan dalam model Random Forest, XGBoost, AdaBoost, dan Extra Trees untuk memprediksi penyakit *liver* dengan lebih akurat.

### C. Pemodelan

Pemodelan dilakukan pada penelitian ini untuk memilih model *ensemble* mana yang terbaik untuk memprediksi keakuratan terhadap penyakit *liver* berdasarkan data pasien. Beberapa evaluasi model yang digunakan pada penelitian ini yaitu:

- Akurasi. Akurasi mengukur persentase prediksi yang benar dibandingkan dengan total jumlah sampel, sehingga memberikan gambaran umum tentang kinerja model dalam mengklasifikasikan data secara keseluruhan [28]. Persamaan Akurasi.

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

- Presisi. Presisi menilai seberapa tepat model dalam memprediksi kelas positif dengan melihat proporsi prediksi positif yang benar, yang sangat penting dalam kasus di mana kesalahan klasifikasi positif dapat berdampak besar [25]. Persamaan Presisi.

$$Presisi = \frac{TP}{TP + FP} \quad (6)$$

- Recall. Recall menunjukkan kemampuan model dalam menangkap seluruh sampel positif dengan benar, di mana nilai yang tinggi menunjukkan bahwa model jarang melewatkan kasus positif yang sebenarnya [28]. Persamaan recall.

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

- F1-Score. F1-Score merupakan rata-rata dari precision dan recall, menjadi metrik yang ideal ketika terdapat ketidakseimbangan data, karena dapat memberikan keseimbangan antara kedua metrik tersebut, memastikan bahwa model tidak hanya fokus pada salah satu aspek [28]. Persamaan F1 Score.

$$F1\ Score = \frac{2 * Presisi * Recall}{Presisi + Recall} \quad (8)$$

- AUC/ROC. AUC/ROC mengukur kemampuan model dalam membedakan antara kelas positif dan negatif dengan

menghitung luas area di bawah kurva ROC, yang memplot *true positive rate* (recall) terhadap *false positive* (FP) rate pada berbagai *threshold* [29]. Persamaan AUC/ROC. Nilai AUC mendekati 1 menunjukkan model yang sangat baik dalam membedakan kelas, sementara nilai mendekati 0.5 menunjukkan performa model kurang baik [25].

Hasil evaluasi dari model yang digunakan pada penelitian ini ditampilkan pada Tabel 2.

Tabel 2. Evaluasi model *ensemble learning*

Model	Accuracy	AUC	Recall	Prec.	F1
<i>Extra Trees</i>	0.7361	0.8082	0.8982	0.7722	0.8293
<i>XG Boost</i>	0.7317	0.7306	0.8619	0.7864	0.8208
<i>Random Forest</i>	0.7253	0.7487	0.8891	0.7668	0.8226
<i>Ada Boost</i>	0.6867	0.6916	0.8258	0.7592	0.7897

Tabel 2 menunjukkan perbandingan performa empat model *ensemble learning* untuk prediksi penyakit *liver*, berdasarkan akurasi, AUC, recall, precision, dan F1-score. *Extra Trees Classifier* (ET) memiliki performa terbaik dengan akurasi tertinggi (73.61%), AUC tertinggi (80.82%), serta recall (89.82%) dan F1-score (82.93%) yang paling baik, menunjukkan bahwa model ini mampu menangkap lebih banyak kasus positif dengan akurasi yang tinggi. XGBoost juga menunjukkan performa yang kuat, terutama dalam precision (78.64%), yang berarti model ini lebih baik dalam mengurangi *false positives*. Random Forest (RF) memiliki performa yang seimbang, tetapi sedikit di bawah ET dan XGBoost. Sementara itu, AdaBoost memiliki performa terendah dalam semua metrik, terutama dalam akurasi (68.67%) dan AUC (69.16%), menunjukkan bahwa model ini kurang optimal dibandingkan metode lainnya. Secara keseluruhan, Extra Trees dan XGBoost merupakan pilihan terbaik untuk model prediksi penyakit *liver* dalam penelitian ini.

#### D. Optimasi Hyperparameter

Pada pemodelan penelitian ini model-model *ensemble learning* digunakan untuk meningkatkan akurasi dari prediksi penyakit *liver* berdasarkan fitur medis pasien dengan menggunakan 10 *cross validation* dengan menggunakan *hyperparameter tuning* dan tanpa menggunakan *hyperparameter*.

Tabel 3. Optimasi model *ensemble*

Model	Akurasi
Random Forest	0.7146
Random Forest + Hyperparameter	0.7253
XGBoost	0.7166
XGBoost + Hyperparameter	0.7317
Adabost	0.7146
Adabost + Hyperparameter	0.6867
Extra Tree	0.7361
Extra Tree + Hyperparameter	0.7361

Dari Tabel 3 dengan menambahkan optimasi *hyperparameter tuning* pada masing-masing model menunjukkan perbandingan akurasi berbagai model *ensemble learning* sebelum dan sesudah *tuning hyperparameter* dalam prediksi penyakit *liver*. Secara umum, *tuning hyperparameter* memberikan peningkatan akurasi pada sebagian besar model. *Random Forest* meningkat dari 71.46% menjadi 72.53%, sedangkan XGBoost mengalami kenaikan dari 71.66% menjadi 73.17%, menunjukkan bahwa model ini mendapatkan manfaat dari optimasi parameter. Sebaliknya, AdaBoost mengalami penurunan dari 71.46% menjadi 68.67%, yang dapat mengindikasikan *overfitting* atau kurangnya parameter yang optimal. Model *Extra Trees* memiliki akurasi tertinggi, yaitu 73.61%, baik sebelum maupun sesudah *tuning*, menunjukkan kestabilan dalam performa. Secara keseluruhan, XGBoost dan *Extra Trees* menjadi model dengan performa terbaik. Sementara itu, *Random Forest* menunjukkan kinerja yang tetap kompetitif setelah dilakukan *tuning hyperparameter*.

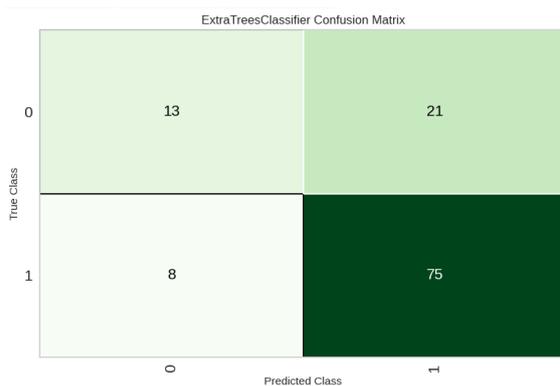
#### E. Analisis Hasil Pemodelan

Berdasarkan hasil evaluasi, *Extra Trees Classifier* memiliki performa terbaik, dengan akurasi tertinggi (73.61%), AUC tertinggi (80.82%), serta F1-score tertinggi (82.93%).

Model ini menunjukkan keseimbangan yang baik antara precision dan recall, sehingga mampu menangkap lebih banyak kasus positif dengan tingkat kesalahan yang rendah. XGBoost juga menunjukkan performa yang kompetitif, terutama dengan precision tertinggi (78.64%), yang mengindikasikan bahwa model ini lebih akurat dalam mengidentifikasi pasien *liver*. *Random Forest* memiliki performa yang cukup baik, tetapi sedikit lebih rendah dibandingkan ET dan XGBoost. AdaBoost memiliki performa terendah, dengan akurasi hanya 68.67% dan AUC 69.16%, yang menunjukkan bahwa model ini kurang optimal dalam menangani dataset yang digunakan.

Hasil ini menunjukkan bahwa *Extra Trees* dan XGBoost lebih unggul dalam menangani prediksi penyakit *liver*, kemungkinan karena kemampuannya dalam menangani variabilitas data dengan baik. *Random Forest* juga memberikan hasil yang cukup stabil, sedangkan AdaBoost tidak bekerja dengan optimal karena lebih rentan terhadap outlier dan kesalahan klasifikasi.

Selanjutnya adalah melihat *confusion matrix* dari model terbaik yang didapatkan berdasarkan hasil penelitian ini yaitu model *Extra Trees* Klasifikasi. *Confusion matrix* digunakan untuk menganalisis performa klasifikasi lebih mendetail, dengan melihat jumlah prediksi yang benar dan salah untuk setiap kelas. Matriks ini terdiri dari *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). *Confusion matrix* ditampilkan pada Gambar 4.

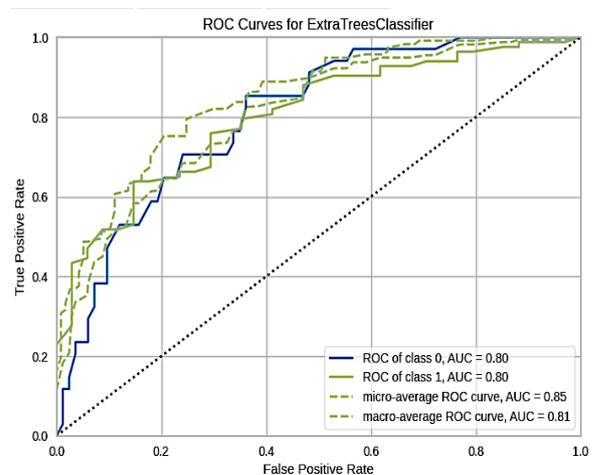


Gambar 4. *Confusion matrix*

Berdasarkan hasil pada gambar 4 Hasil *confusion matrix* pada model *Extra Trees Classifier* menunjukkan bahwa model ini mampu mengklasifikasikan 75 sampel pasien *liver* dengan benar (*True Positives*), sementara terdapat 8 kasus pasien *liver* yang salah diklasifikasikan sebagai non-pasien (*False Negatives*). Hal ini mengindikasikan bahwa model memiliki recall yang tinggi, yang berarti sebagian besar pasien *liver* berhasil teridentifikasi dengan baik. Di sisi lain, model juga berhasil mengklasifikasikan 13 sampel non-pasien *liver* dengan benar (*True Negatives*), tetapi terdapat 21 kasus non-pasien yang salah diklasifikasikan sebagai pasien *liver* (*False Positives*). Kesalahan ini menunjukkan bahwa model cenderung lebih sensitif terhadap kasus positif, yang menyebabkan beberapa non-pasien *liver* dikategorikan sebagai pasien *liver*.

Secara keseluruhan, model memiliki kemampuan prediksi yang cukup baik, dengan keseimbangan antara recall yang tinggi dan precision yang cukup baik, meskipun masih terdapat *false positives* yang dapat dikurangi melalui penyesuaian *threshold* prediksi atau optimasi parameter model.

Selain melihat hasil *confusion matrix* pada pembahasan ini juga ditampilkan kurva AUC/ROC dari model terpilih pada Gambar 5.



Gambar 5. Kurva AUC/ROC

Pada gambar 5 model *Extra Trees Classifier* memiliki AUC sebesar 0.80 untuk masing-masing kelas, serta *macro-average* AUC 0.81

dan *micro-average* AUC 0.85. Nilai ini menunjukkan bahwa model cukup baik dalam memprediksi kedua kelas, dengan kemampuan membedakan pasien *liver* yang cukup tinggi.

#### V. KESIMPULAN

Penelitian ini telah mengevaluasi performa beberapa model *ensemble learning* yaitu *Random Forest*, *XGBoost*, *AdaBoost* dan *Extra Trees Classifier* dalam memprediksi penyakit *liver* berdasarkan data pasien. Dari hasil eksperimen, *Extra Trees Classifier* terbukti menjadi model terbaik, dengan akurasi tertinggi (73.61%) dan AUC 76.82%, serta memiliki keseimbangan optimal antara recall (89.82%) dan precision (77.22%). Model ini memiliki kemampuan tinggi dalam mengidentifikasi pasien *liver* dan memberikan prediksi yang cukup andal berdasarkan analisis *confusion matrix* dan kurva ROC. Meskipun performa model sudah cukup baik, masih terdapat beberapa kesalahan klasifikasi, terutama pada pasien non-*liver* yang diklasifikasikan sebagai pasien *liver* (*false positives*). Untuk itu, saran untuk penelitian selanjutnya model ini masih dapat ditingkatkan dengan teknik *balancing* data, tuning *hyperparameter* yang lebih optimal, serta penggunaan dataset yang lebih besar dan bervariasi.

#### DAFTAR PUSTAKA

- [1] M. Sharma and R. Parveen, "A comparative study of data mining, digital image processing and genetical approach for early detection of *liver* cancer," *2020 IEEE International Conference on Computing, Power and Communication Technologies, GUCON 2020*, pp. 687–692, 2020, doi: 10.1109/GUCON48875.2020.9231193.
- [2] A. P. Ayudhitama and Utomo Pujianto, "Analisa 4 Algoritma Dalam Klasifikasi *Liver* Menggunakan Rapidminer," *Jurnal Informatika Polinema*, vol. 6, no. 2, pp. 1–9, 2020, doi: 10.33795/jip.v6i2.274.
- [3] F. L. D. Cahyanti, F. Sarasati, W. Astuti, and E. Firasari, "Klasifikasi Data Mining Dengan Algoritma Machine Learning Untuk Prediksi Penyakit *Liver*," *Technologia : Jurnal Ilmiah*, vol. 14, no. 2, p. 134, 2023, doi: 10.31602/tji.v14i2.10093.
- [4] M. Kesuma, "Prediksi Penyakit *Liver* Menggunakan Algoritma Random Forest," *Jurnal informasi dan Komputer*, vol. 11, no. 2, p. 2023, 2023.
- [5] P. Tripathi *et al.*, "Classification of nucleic acid amplification on ISFET arrays using spectrogram-based neural networks," *Computers in Biology and Medicine*, vol. 161, no. October 2022, p. 107027, 2023, doi: 10.1016/j.compbio.2023.107027.
- [6] M. Nurkholifah, Jasmarizal, Y. Umar, and Rahmaddeni, "Analisa Performa Algoritma Machine Learning Dalam Prediksi Penyakit *Liver*," *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, vol. 4, no. 1, pp. 164–172, 2023, doi: 10.35870/jimik.v4i1.149.
- [7] A. Sahu, M. Gaur, N. C. Mahanandia, E. Subudhi, R. P. Swain, and B. B. Subudhi, "Identification of core therapeutic targets for Monkeypox virus and repurposing potential of drugs against them: An in silico approach," *Computers in Biology and Medicine*, vol. 161, no. October 2022, p. 106971, 2023, doi: 10.1016/j.compbio.2023.106971.
- [8] I. R. Hikmah and R. N. Yasa, "Perbandingan Hasil Prediksi Diagnosis pada Indian *Liver* Patient Dataset (ILPD) dengan Teknik Supervised Learning Menggunakan Software Orange," *Jurnal Telematika*, vol. 16, no. 2, pp. 69–76, 2022, doi: 10.61769/telematika.v16i2.402.
- [9] I. N. Abrar, A. Abdullah, and S. Sucipto, "*Liver* Disease Classification Using the Elbow Method to Determine Optimal K in the K-Nearest Neighbor (K-NN) Algorithm," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 2, pp. 218–228, 2023, doi: 10.32736/sisfokom.v12i2.1643.
- [10] S. R. Velu, V. Ravi, and K. Tabianan, "Data mining in predicting *liver* patients using classification model," *Health and Technology*, vol. 12, no. 6, pp. 1211–1235, 2022, doi: 10.1007/s12553-022-00713-3.
- [11] S. Zulaikhah Hariyanti Rukmana, A. Aziz, and W. Harianto, "Optimasi Algoritma K-Nearest Neighbor (Knn) Dengan Normalisasi Dan Seleksi Fitur Untuk Klasifikasi Penyakit *Liver*," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 439–445, 2022, doi: 10.36040/jati.v6i2.4722.
- [12] U. Indahyanti, N. L. Azizah, and H. Setiawan, "Pendekatan *Ensemble learning* Untuk Meningkatkan Akurasi Prediksi Kinerja Akademik Mahasiswa," *Jurnal Sains dan Informatika*, vol. 8, no. 2, pp. 160–169, 2022, doi: 10.34128/jsi.v8i2.459.
- [13] A. H. Yusufi, A. Kharisma, A. D. Adinata, D. F. Ramzy, and M. M. Santoni, "Prediksi Resiko Kematian Pada Penderita Penyakit Kardiovaskular Menggunakan Metode *Ensemble learning*," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 531–542, 2022.
- [14] Y. Wang *et al.*, "The theranostic value of acetylation gene signatures in obstructive sleep

- apnea derived by machine learning,” *Computers in Biology and Medicine*, vol. 161, no. February, 2023, doi: 10.1016/j.compbimed.2023.107058.
- [15] P. Mahajan, S. Uddin, F. Hajati, and M. A. Moni, “Ensemble learning for Disease Prediction: A Review,” *Healthcare (Switzerland)*, vol. 11, no. 12, 2023, doi: 10.3390/healthcare11121808.
- [16] N. Nahar, F. Ara, M. A. I. Nelay, V. Barua, M. S. Hossain, and K. Andersson, “A Comparative Analysis of the Ensemble Method for Liver Disease Prediction,” *ICIET 2019 - 2nd International Conference on Innovation in Engineering and Technology*, pp. 23–24, 2019, doi: 10.1109/ICIET48527.2019.9290507.
- [17] M. A. Hasanah, S. Soim, and A. S. Handayani, “Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir,” *Journal of Applied Informatics and Computing*, vol. 5, no. 2, pp. 103–108, 2021, doi: 10.30871/jaic.v5i2.3200.
- [18] S. Navisa, Luqman Hakim, and Aulia Nabilah, “Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM,” *Jurnal Sistem Cerdas*, vol. 4, no. 2, pp. 114–125, 2021, doi: 10.37396/jsc.v4i2.162.
- [19] Y. A. Singgalen, “Analisis Performa Algoritma NBC, DT, SVM dalam Klasifikasi Data Ulasan Pengunjung Candi Borobudur Berbasis CRISP-DM,” *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 3, pp. 1634–1646, 2022, doi: 10.47065/bits.v4i3.2766.
- [20] Y. A. Azis, “Diagram Alir Penelitian: Pengertian, Jenis dan Contoh,” Online. [Online]. Available: <https://deepublishstore.com/>
- [21] A. Ishaq *et al.*, “Improving the Prediction of Heart Failure Patients’ Survival Using SMOTE and Effective Data Mining Techniques,” *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.
- [22] P. A. Moreno-Sanchez, “Development of an Explainable Prediction Model of Heart Failure Survival by Using Ensemble Trees,” *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 4902–4910, 2020, doi: 10.1109/BigData50022.2020.9378460.
- [23] S. Elina, H. Yulianti, O. Soesanto, and Y. Sukmawaty, “Penerapan Metode Extreme Gradient Boosting ( XGBOOST ) pada Klasifikasi Nasabah Kartu Kredit,” *Journal of Mathematics: Theory and Applications*, vol. 4, no. 1, pp. 21–26, 2022.
- [24] Andreyestha and A. Subekti, “Analisis Sentiment pada Ulasan Film Dengan Optimasi Ensemble learning,” *JURNAL INFORMATIKA*, vol. 7, no. 1, pp. 5–8, 2020.
- [25] D. S. Ali and M. A. Aljabery, “Predicting Liver Cirrhosis Stages Using Extra Trees, Random Forest, and SVM with Data Mining Techniques,” *Informatica (Slovenia)*, vol. 48, no. 21, pp. 15–26, 2024, doi: 10.31449/inf.v48i21.6752.
- [26] JeevanNagaraj, “Indian Liver Patient Dataset,” Online. [Online]. Available: <https://www.kaggle.com/datasets/jeevannagaraj/indian-liver-patient-dataset/data>
- [27] H. M. Nawawi, M. Iqbal, Y. Yudhistira, I. Nawawi, S. Widodo, and N. Herlinawati, “Deep learning for Tesla’s stock prices prediction,” *AIP Conference Proceedings*, vol. 2714, no. October 2022, 2023, doi: 10.1063/5.0128535.
- [28] A. Mustopa *et al.*, “Ekstraksi Fitur dengan Classifier Random Forest untuk Memprediksi Covid 19 Berdasarkan Hasil Rontgen Thorax Feature Extraction with Random Forest Classifier to Predict Covid 19 Based on Results Thorax X - ray,” *Jurnal Sistem Informasi*, vol. 11, pp. 515–525, 2022.
- [29] E. Edric and S. P. Tamba, “Prediksi Penyakit Gagal Jantung Dengan Menggunakan Random Forest,” *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 176–181, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2445.