

Evaluasi Model *Machine Learning* untuk Klasifikasi Potensi Keberadaan Ikan Kembung (*Rastrelliger sp.*) di Selat Sunda

Putri Adi Khotimah*, Willdan Aprizal Arifin, Ayang Armelita Rosalia

Departemen Sistem Informasi Kelautan, Kampus Serang, Universitas Pendidikan Indonesia

Jl. Dr. Setiabudhi No. 229, Kota Bandung, Jawa Barat, 40154 Indonesia

Corresponding author, e-mail: putriadikhotimah@upi.edu

ABSTRAK: Karakteristik lingkungan perairan Selat Sunda berperan penting dalam menentukan keberadaan ikan kembung. Variabilitas parameter lingkungan menyebabkan pola keberadaan ikan kembung menjadi dinamis sehingga perlu adanya pendekatan yang mampu menangkap kompleksitas hubungan antar variabel salah satunya menggunakan machine learning. Penelitian ini bertujuan untuk membuat pemodelan klasifikasi keberadaan ikan kembung menggunakan algoritma Random Forest dan XGBoost dengan pendekatan time series untuk mempertimbangkan aspek temporal data. Proses pelatihan model menggunakan data tahun 2019-2023 dan pengujian model menggunakan data tahun 2024-2025. Hyperparameter tuning dilakukan untuk mengoptimalkan kinerja model. Hasil evaluasi, Random Forest menunjukkan performa yang lebih unggul dengan akurasi 95% sementara XGBoost memperoleh akurasi sebesar 94%. Analisis menggunakan pendekatan SHAP menunjukkan bahwa klorofil-a menjadi variabel yang paling dominan dalam menentukan prediksi keberadaan ikan kembung di Selat Sunda. Pola musiman menunjukkan bahwa puncak keberadaan ikan kembung berada di periode Mei-Juni dan musim penurunan keberadaan pada periode November-Februari.

Kata kunci: Keberadaan Ikan Kembung; Pola Musiman; Random Forest; XGBoost

*Evaluation of Machine Learning Models for Classifying the Potential Presence of Mackerel (*Rastrelliger sp.*) in the Sunda Strait*

ABSTRACT: The environmental characteristics of the Sunda Strait play a crucial role in determining the presence of mackerel. Variability in environmental parameters causes mackerel distribution patterns to be dynamic, necessitating an approach capable of capturing the complexity of relationships between variables—one such approach is machine learning. This study aims to develop a classification model for mackerel presence using the Random Forest and XGBoost algorithms, with a time-series approach to account for the temporal aspects of the data. Model training utilized data from 2019–2023, and model testing used data from 2024–2025. Hyperparameter tuning was performed to optimize model performance. Evaluation results show that Random Forest demonstrated superior performance with 95% accuracy, while XGBoost achieved 94% accuracy. Analysis using the SHAP approach indicates that chlorophyll-a is the most dominant variable in predicting mackerel presence in the Sunda Strait. Seasonal patterns indicate that the highest presence of mackerel occurs during the May–June period, with a lean season during November–February.

Keywords: Mackerel Presence; Seasonal pattern; Random Forest; XGBoost.

PENDAHULUAN

Selat Sunda terletak di antara Pulau Jawa dan Sumatera, merupakan jalur penghubung antara Samudra Hindia dan Laut Jawa. Dari sisi oseanografi, perairan Selat Sunda berperan penting dalam sirkulasi massa air di Indonesia, yang dinamikanya dipengaruhi oleh massa air Laut Jawa

dan Samudra Hindia. Perairan di Selat Sunda dipengaruhi oleh pola angin muson barat, peralihan dan muson timur (Firdaus *et al.*, 2018). Produktivitas perikanan di perairan ini relatif tinggi, terutama untuk jenis ikan pelagis kecil. Salah satu komoditas pelagis kecil yang penting di wilayah Selat Sunda adalah ikan kembung. Ikan kembung termasuk salah satu ikan yang dominan tertangkap di Selat Sunda dengan total tangkapan 233 ton per tahun (Tarigan *et al.*, 2025). Ikan kembung juga berperan penting dalam memenuhi kebutuhan protein masyarakat pesisir, sehingga meningkatkan permintaan pasar dan aktivitas penangkapan. Musim penangkapan ikan kembung saat ini cenderung tidak menentu dan sulit diprediksi yang berdampak langsung pada pendapatan nelayan, sehingga perlu dilakukan pengkajian untuk menyediakan informasi tambahan sebagai dasar untuk merancang strategi pengembangan pengelolaan perikanan ikan kembung di Selat Sunda (Caesario *et al.*, 2022).

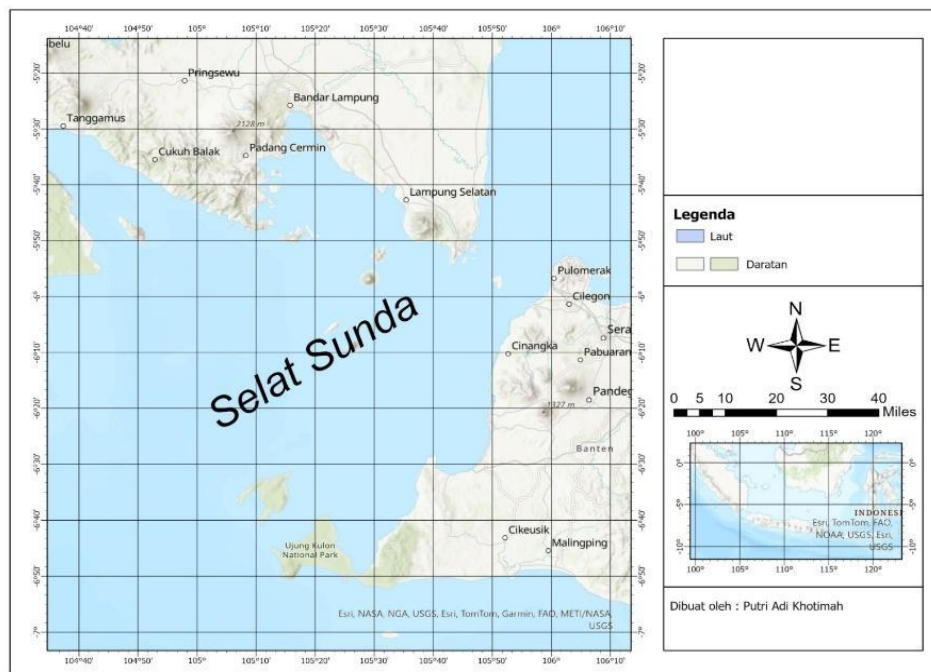
Karakteristik fisik-kimia perairan dan perubahan musiman parameter oseanografi sangat memengaruhi keberadaan ikan kembung di Selat Sunda. Faktor lingkungan memegang peranan penting dalam mencegah gangguan pada stok populasi ikan kembung (Dawolo *et al.*, 2025). Iklim, kedalaman termoklin, ketersediaan makanan, klorofil-a, suhu, salinitas, oksigen, siklus massa air dan kecepatan arus merupakan faktor lingkungan untuk mengetahui keberadaan ikan di suatu perairan (Dwiyanti *et al.*, 2022). Variabel-variabel tersebut memiliki pola temporal yang penting bagi keberadaan ikan kembung. Indikator SPL dan klorofil-a yang berimplikasi langsung pada distribusi ikan pelagis, serta dapat mengidentifikasi fenomena *upwelling* di Selat Sunda (Ramadlanie *et al.*, 2022). Klorofil-a sebagai indikator produktivitas primer perairan berkorelasi positif dengan hasil tangkapan ikan kembung (Titaheluw *et al.*, 2022). Salinitas juga berperan penting dalam menentukan distribusi ikan kembung serta mendukung proses pemijahan. Ikan pelagis kecil, termasuk ikan kembung, memiliki kemampuan terbatas dalam mentoleransi faktor lingkungan seperti salinitas (Asbar & Ihsan, 2022). Salinitas merupakan variabel yang paling berpengaruh positif terhadap CPUE ikan pelagis kecil, termasuk ikan kembung (Puspasari *et al.*, 2021). Selain itu, curah hujan berpengaruh dalam menurunkan salinitas dan memengaruhi produktivitas primer di perairan. SPL dan curah hujan memengaruhi kelimpahan ikan tongkol dan ikan pelagis kecil (Lubis *et al.*, 2024).

Variabilitas parameter oseanografi menyebabkan pola sebaran dan potensi keberadaan ikan kembung menjadi dinamis dan sulit diprediksi, sehingga diperlukan pendekatan yang mampu menangkap kompleksitas hubungan antarvariabel lingkungan secara lebih adaptif untuk mendukung penangkapan ikan yang optimal dan berkelanjutan. Salah satu pendekatan yang bisa dilakukan adalah membuat pemodelan menggunakan *machine learning*, khususnya untuk memodelkan klasifikasi potensi keberadaan ikan kembung. Algoritma *machine learning* yang sering digunakan untuk model klasifikasi adalah Random Forest dan XGBoost. Random Forest merupakan salah satu metode klasifikasi terbaik dan banyak digunakan dalam *machine learning* (Rizky & Andarsyah, 2023). XGBoost memiliki kemampuan yang baik dalam menangani data yang kompleks dan besar secara efektif, serta menghasilkan prediksi yang akurat untuk klasifikasi (Sihombing *et al.*, 2025). Penelitian Kurnianto *et al.* (2024) membandingkan Random Forest dan SVM untuk klasifikasi daerah penangkapan ikan menggunakan parameter oseanografi, penelitian ini belum menganalisis pola temporal musiman dan tidak spesifik pada ikan pelagis kecil seperti ikan kembung. Penelitian lainnya yang menggunakan *machine learning* dengan algoritma K-NN, parameter yang digunakan SPL dan curah hujan untuk melakukan klasifikasi keberadaan ikan tongkol di Laut Jawa, belum menambahkan parameter lain seperti klorofil-a dan salinitas dan belum membandingkan dengan algoritma lain (Lubis *et al.*, 2024). Penelitian pemodelan distribusi atau keberadaan ikan demersal di perairan Aceh, membandingkan XGBoost dan SVR, menunjukkan XGBoost memberikan performa lebih baik, penelitian ini menggunakan pendekatan regresi dan belum mengkaji spesies pelagis kecil yang bergantung pada variabilitas oseanografi (Purnawan *et al.*, 2025). Namun, meskipun sudah ada penelitian yang menggunakan algoritma *machine learning* untuk analisis keberadaan ikan, kajian yang secara khusus membandingkan Random Forest dan XGBoost dalam pemodelan potensi keberadaan ikan kembung di perairan Selat Sunda masih sangat terbatas. Oleh karena itu, penelitian ini bertujuan untuk membangun model klasifikasi keberadaan ikan kembung menggunakan kedua algoritma tersebut serta menganalisis interaksi faktor lingkungan yang memengaruhi potensi keberadaan ikan kembung.

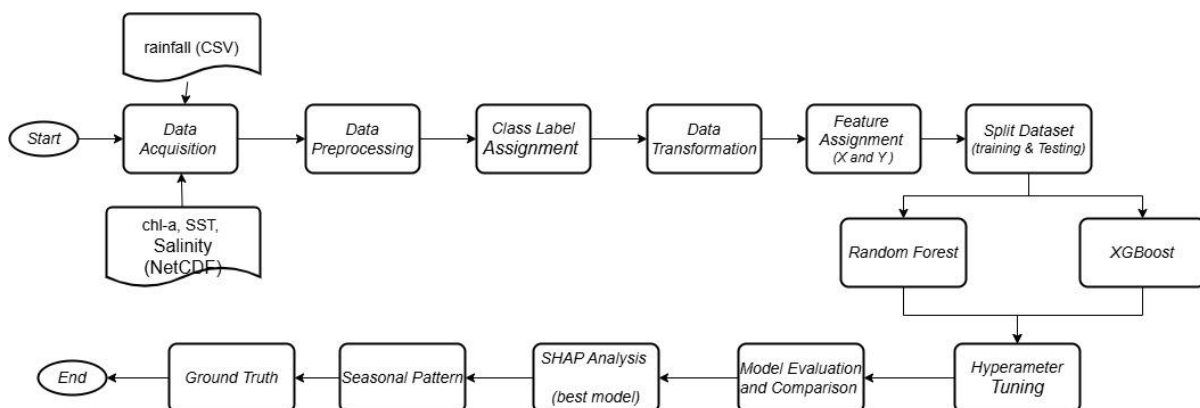
MATERI DAN METODE

Perairan Selat Sunda berada pada koordinat $5^{\circ}55'8.4''$ LS dan $105^{\circ}53'5.4''$ BT. Lokasi ini dipilih karena merupakan salah satu daerah penangkapan ikan pelagis kecil khususnya ikan kembung yang penting di Indonesia. Ikan kembung di Selat Sunda menyumbang 11% dari total tangkapan ikan pelagis kecil (Tarigan *et al.*, 2025). Penelitian dilakukan pada bulan Februari 2026 hingga Mei 2026 dengan menggunakan pendekatan kuantitatif dan didukung validasi lapangan. *Software* yang digunakan untuk pengolahan data adalah Google Colab.

Berdasarkan *flowchart* pada gambar 2, penelitian ini diawali dengan pengumpulan data atau *data acquisition* parameter oseanografi. Parameter yang digunakan adalah suhu permukaan laut (SPL), salinitas, klorofil-a dan curah hujan. Data SPL dan salinitas bersumber dari CMEMS (*Copernicus Marine Environment Monitoring Service*) dengan resolusi horizontal $1/12^{\circ}$ (~8km) dan resolusi temporal harian dengan format NetCDF. Untuk data klorofil-a bersumber dari Copernicus



Gambar 1. Layout Lokasi Penelitian



Gambar 2. Flowchart Penelitian

CMEMS, resolusi $1/4^{\circ}$ (~25km) dan resolusi temporal harian dengan format NetCDF. Data curah hujan dikumpulkan melalui *website* <https://climateengine.org> menggunakan produk CHIRPS dengan resolusi spasial 4.8km dan resolusi temporal harian. Data harian yang terkumpul selama tujuh tahun (2019-2025). *Data preprocessing*, berupa ekstraksi file yang berformat NetCDF menjadi data tabular dengan metode ekstraksi *data tergrid* menggunakan *library xarray* pada pemrograman bahasa *python*. *filter* wilayah dilakukan dengan *spatial masking* menggunakan *geopandas* dan *shapely* untuk memisahkan hanya wilayah perairan Selat Sunda. *Data cleaning* menggunakan metode *listwise deletion*, yaitu menghapus data yang mengandung nilai kosong.

Selain data sekunder dari citra satelit, metode wawancara juga dilakukan bersama nelayan lokal di Karangantu dan pesisir pantai Lelean yang melakukan *trip* hingga ke Selat Sunda. Data wawancara sebagai variabel pendukung digunakan secara deskriptif untuk validasi temporal dan konfirmasi pola musiman hasil prediksi model selaras dengan yang dilaporkan oleh nelayan.

Kelas target ditentukan berdasarkan kriteria SPL, salinitas, dan klorofil-a yang optimum untuk mendukung keberadaan ikan kembung. Menurut penelitian Tarigan *et al.* (2020), konsentrasi optimum klorofil-a untuk ikan kembung adalah $>0,2 \text{ mg/m}^3$. Sementara untuk suhu optimum ikan kembung berkisar antara $29 - 31^{\circ}\text{C}$ (Purba *et al.*, 2025). Salinitas optimum untuk ikan kembung berkisar antara $< 33 \text{ psu}$ (Apriansyah & Utami, 2024).

Transformasi data bertujuan untuk meningkatkan kinerja model. Pada variabel bulan dilakukan pendekatan trigonometri *Cyclical Encoding* untuk menangkap pola musiman yang masuk dalam kategorikal siklik. Kategorikal siklik dimanfaatkan *machine learning* untuk meningkatkan kemampuan model dan menangkap data temporal (Luangmaneerote & Thasana, 2025). Rumus persamaan untuk trigonometri siklik ditulis menurut Rashid *et al.*, (2025) :

$$x_{\sin} = \sin\left(\frac{2\pi * \text{month}}{12}\right)$$

$$x_{\cos} = \cos\left(\frac{2\pi * \text{month}}{12}\right)$$

Variabel SPL dan klorofil-a ditransformasikan menggunakan pendekatan *lagging*. Pendekatan ini didasarkan pada mekanisme rantai makanan menunjukkan bahwa produktivitas *primer* periode sebelumnya memengaruhi distribusi ikan kembung pada periode berikutnya. Ketergantungan lagged antara SPL dan klorofil-a dalam model statistik menunjukkan pentingnya efek jeda (*lag effect*) dalam model di masa depan (Dunstan *et al.*, 2018). Rumusnya ditulis dengan persamaan: $t - 1$

Data labeling atau pelabelan data yaitu label X (independen) yang terdiri dari data bulan, SPL, salinitas, klorofil-a dan curah hujan. Label Y (dependen) yang merupakan kelas target yaitu kelas potensi keberadaan ikan kembung. Data yang telah diberi label kemudian melalui proses *data splitting*, data dibagi secara temporal (*time series*) tanpa pengacakan dengan tujuan untuk mencegah *data leakage*. Pendekatan metode *time series* dalam pemodelan *machine learning* terbukti menghasilkan kinerja terbaik (Prasetyo *et al.*, 2024). Data dibagi menjadi 2 yaitu *training set* dan *testing set* dengan proporsi pembagian data adalah 70:30. Pendekatan ini dipilih untuk menjaga urutan temporal data dan meningkatkan akurasi prediksi pada data yang belum diamati. Metode ini bertujuan untuk menghindari kebocoran data yang dapat memengaruhi validasi model.

Tabel 1. Klasifikasi Kelas Potensi

Parameter	Kelas Potensi	
	0	1
Suhu ($^{\circ}\text{C}$)	< 29 atau > 31	$29 - 31$
Salinitas (psu)	> 33	< 33
Klorofil-a (mg/m^3)	< 0.2	> 0.2

Tabel 2. Pembagian Data

<i>Subset</i>	<i>Periode</i>	Jumlah Data
<i>Training</i>	1 Jan 2019 – 31 Des 2023	1825
<i>Testing</i>	1 Jan 2024 – 31 Des 2025	731
<i>Total</i>	2019 - 2025	2556

Tahap berikutnya yaitu pemodelan, merupakan tahapan utama dalam penelitian. Metode yang dipakai untuk membuat model klasifikasi adalah Random Forest dan XGBoost dengan menggunakan pendekatan *time series*. Random Forest merupakan algoritma *ensemble* berbasis *bagging* yang membangun sejumlah pohon keputusan dengan mekanisme *voting* mayoritas. Algoritma Random Forest memberikan kinerja kompetitif dibandingkan algoritma klasik (Agustina *et al.*, 2025). Namun, kinerja algoritma Random Forest dapat dipengaruhi oleh beberapa parameter. Penggunaan pohon yang terlalu sedikit ataupun terlalu banyak dapat melemahkan kinerja pohon atau *decision tree* (Fauzi *et al.*, 2025). Metode Random Forest unggul karena memiliki kemampuan dalam menangani data yang kompleks dan mampu memberikan hasil yang baik (Fadli & Saputra, 2023). Pada algoritma Random Forest ketidakseimbangan kelas ditangani dengan parameter *class_weigh* = '*balanced_subsample*'. Algoritma Random Forest dimulai dengan menentukan *subset* data menggunakan teknik *bootstrap aggregating* atau *bagging*, kemudian setelah *subset* terbentuk Random Forest akan membangun *decision tree* secara *independen* untuk setiap *sample bootstrap*. Pada setiap simpul, pemilihan fitur terbaik dalam Random Forest menggunakan metrik *Gini impurity* untuk pemisahan data yang optimal yaitu p_i adalah probabilitas memilih item dengan kelas i dari kumpulan data, dan c adalah jumlah total kelas. Rumus prediksi akhir untuk klasifikasi dan gini dituliskan dalam persamaan berikut (Ernawati *et al.*, 2025):

$$Gini(s) = 1 - \sum_{i=1}^c p_i^2$$

XGBoost merupakan algoritma *ensemble* berbasis *boosting* yang membangun pohon secara bertahap, setiap pohon memperbaiki kesalahan prediksi pohon sebelumnya. Setiap pohon dalam *ensemble* ini berkontribusi dalam mengurangi error menggunakan pendekatan *boosting* (Yulianti *et al.*, 2022). XGBoost menerapkan regulasi yang mengontrol kompleksitas model guna mencegah *overfitting*, serta menggunakan teknik *subsampling* untuk meningkatkan kemampuan generalisasi. Meskipun algoritma ini mampu menangani data yang kompleks, kualitas prediksi tetap sangat bergantung pada kualitas data *input* sehingga kesalahan pada data dapat memengaruhi hasil model. Ketidakseimbangan kelas pada model XGBoost ditangani dengan parameter *scale_pos_weight*. Fungsi objektif yang digunakan adalah *logistic loss* (*log loss*) untuk klasifikasi biner. Adapun rumus regulasi ditulis dalam persamaan berikut (Illahiyah *et al.*, 2026) :

$$Obj = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^k \Omega(f_k)$$

Optimasi model dilakukan melalui dua tahapan, yaitu hyperparameter search melalui pendekatan Randomized Search sebagai eksplorasi awal dan Grid Search sebagai fine-tuning yang berfokus pada pencarian di sekitar nilai terbaik pada tahap pertama. Kedua tahapan menggunakan *Time Series Cross Validation* dengan 10 *fold* untuk mempertahankan urutan temporal data dan menghindari kebocoran data (Filiz *et al.*, 2025). Evaluasi model menggunakan nilai evaluasi *Confusion matrix* yang terdiri dari data aktual sehingga diketahui nilai *accuracy*, *precision*, *recall*, dan *F1-Score* (Kurnianto *et al.*, 2024). Nilai aktual tersebut dapat dituliskan dalam persamaan rumus sebagai berikut (Kurnianto *et al.*, 2024).

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

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

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

$$F1\ score = 2 \frac{precision * recall}{precision + recall}$$

Matriks terdiri dari empat komponen yaitu *True Positive*, *True Negative*, *False Positive* dan *False Negative*. Matriks evaluasi precision yaitu proporsi prediksi yang benar dari keseluruhan data, nilai ini menunjukkan seberapa banyak prediksi positif yang benar. *Recall* untuk mengukur kemampuan model dalam mendeteksi seluruh data positif, *F1-Score* yaitu rata-rata antara nilai precision dan *recall*. *Confusion matrix* digunakan karena kemampuannya dalam memberikan gambaran menyeluruh terhadap distribusi kesalahan model secara terperinci, yang tidak dapat diperoleh dari nilai akurasi tunggal lainnya (Valerian *et al.*, 2025).

HASIL DAN PEMBAHASAN

Pada gambar 3, menunjukkan dinamika musiman yang signifikan di perairan Selat Sunda pada periode 2019 hingga 2025. SPL di Selat Sunda berada di puncak tertinggi pada bulan Mei hingga Juni $\pm 30.5^{\circ}\text{C}$ yang bertepatan dengan musim peralihan I dan suhu terendah pada bulan Agustus hingga September $\pm 29.3^{\circ}\text{C}$ yang menandakan pengaruh *upwelling* musim timur. Standar deviasi menunjukkan kondisi termal perairan Selat Sunda relatif stabil antartahun. Klorofil-a menunjukkan pola yang berlawanan dengan SPL, konsentrasi signifikan sejak bulan Agustus dan mencapai puncaknya pada Oktober dan November $\pm 0.4\ \text{mg/m}^3$. Salinitas menunjukkan fluktuasi yang berkaitan dengan siklus penguapan dan curah hujan. Bulan Maret menunjukkan nilai salinitas terendah $\pm 31.85\ \text{psu}$ bertepatan dengan akhir musim barat. Puncak salinitas tertinggi pada akhir tahun pada bulan November hingga Desember $\pm 32.90\ \text{psu}$. Pola curah hujan di Selat Sunda bertipe monsunial dengan intensitas terendah terjadi pada bulan Juli hingga Agustus yang menandakan puncak musim kemarau. Peningkatan drastis mulai pada bulan September dan mencapai puncaknya di bulan November. Hubungan peningkatan signifikan curah hujan dan klorofil-a di akhir tahun menunjukkan masukan air tawar dari daratan berperan penting dalam membawa nutrisi ke perairan pantai Selat Sunda. Curah hujan dan klorofil-a menunjukkan korelasi yang positif, menunjukkan klorofil-a tinggi saat curah hujan tinggi, dan sebaliknya klorofil-a rendah saat curah hujan rendah (Zulfa *et al.*, 2024).

Transformasi Variabel Model

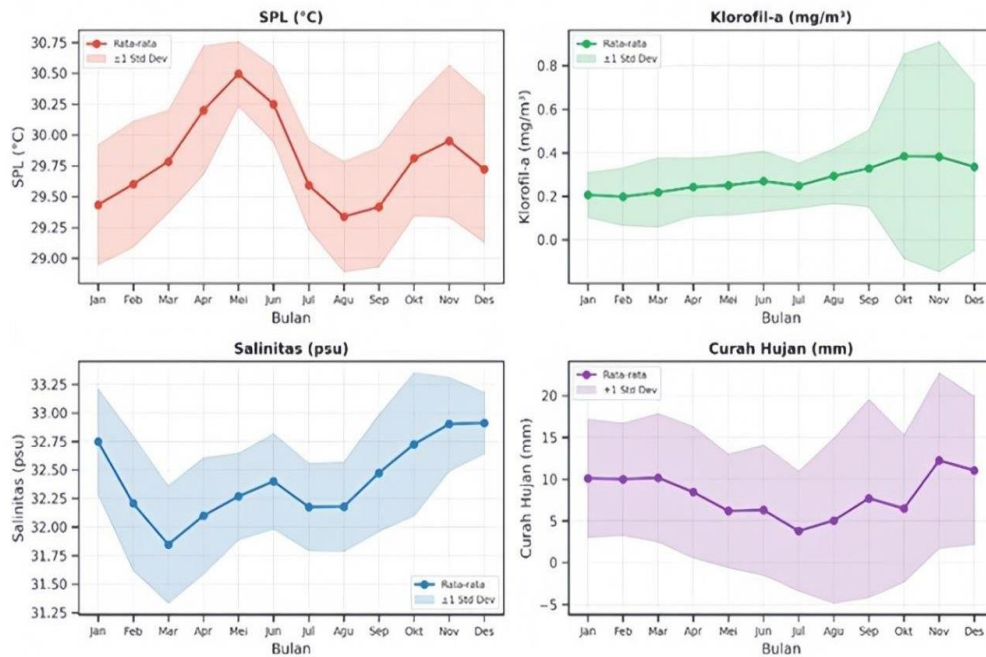
Transformasi untuk variabel SPL dan klorofil-a menggunakan kondisi *Lag1* atau nilai periode sebelumnya. Ikan pelagis kecil menunjukkan respon yang sensitif terhadap fluktuasi lingkungan laut yang temporal seperti SPL dan ketersediaan pakan (Balde *et al.*, 2019). Kondisi oseanografi pada $t-1$ terbukti menjadi prediktor dominan dalam pemodelan parameter laut (Durap, 2025). Salinitas tidak ditransformasi oleh *Lag1* karena sifatnya relatif persisten. Menurut Alkita *et al.* (2022), distribusi ikan pelagis kecil di perairan Indonesia menunjukkan respons terhadap salinitas secara langsung, menunjukkan bahwa ikan pelagis kecil termasuk ikan kembung bereaksi terhadap kondisi salinitas aktual, bukan historisnya.

Variabel Bulan yang merupakan kategorial siklik ditransformasikan menggunakan fitur trigonometri *Cyclical Encoding*. Tujuan dari transformasi ini agar model dapat menangkap bahwa bulan 12 (Desember) dan bulan 1 (Januari) memiliki kedekatan temporal. Hasil akhir transformasi

fitur dapat dilihat pada gambar 4, kemudian sebaran data secara keseluruhan menggunakan visualisasi boxplot dapat dilihat pada gambar 5.

Pengembangan Model

Pemodelan menggunakan dua algoritma *machine learning* yaitu Random Forest dan XGBoost untuk membandingkan performa kedua algoritma dengan hasil kinerja yang optimal. Optimasi model dengan pendekatan *hyperparameter tuning*, diharapkan dapat memperoleh



Gambar 3. Grafik rata-rata bulanan Oseanografi di Selat Sunda 2019-2025

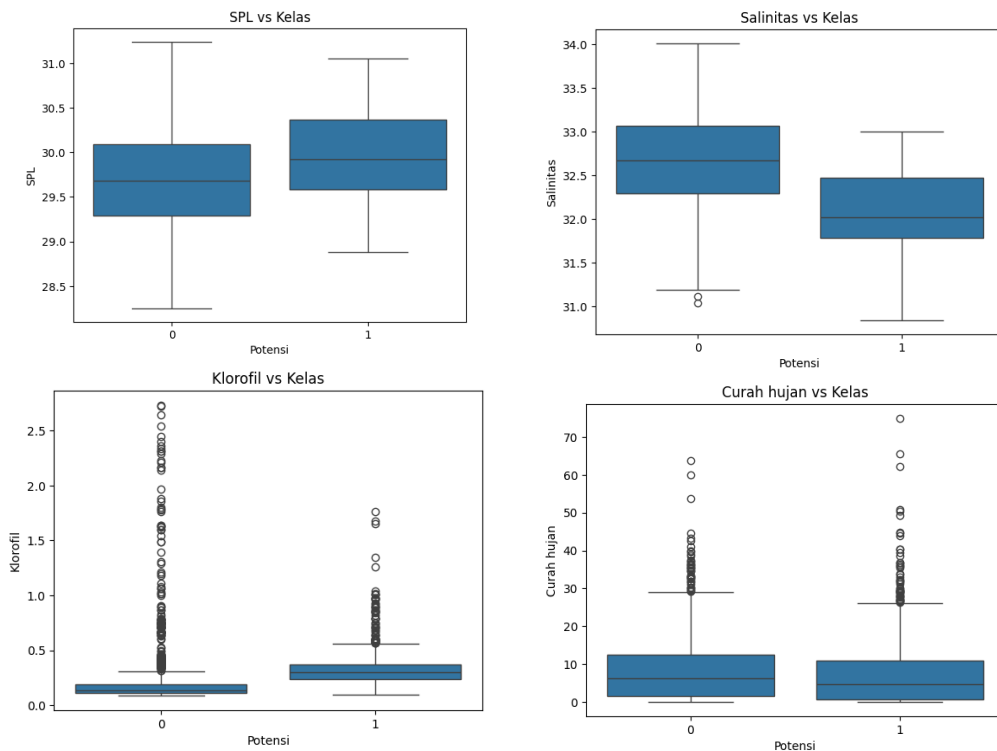
	Bulan_sin	Bulan_cos	SPL	Salinitas	Klorofil	Curah hujan	Potensi
0	5.000000e-01	0.866025	28.77	33.19	0.10	13.07	0
1	5.000000e-01	0.866025	28.75	33.20	0.11	1.98	0
2	5.000000e-01	0.866025	28.71	33.22	0.11	0.00	0
3	5.000000e-01	0.866025	28.81	33.21	0.11	1.34	0
4	5.000000e-01	0.866025	29.06	33.16	0.10	0.01	0
...	...	...	...	...	...	...	...
2551	-2.449294e-16	1.000000	29.19	33.28	0.17	5.85	0
2552	-2.449294e-16	1.000000	29.12	33.28	0.17	9.38	0
2553	-2.449294e-16	1.000000	29.12	33.26	0.17	17.47	0
2554	-2.449294e-16	1.000000	29.11	33.26	0.17	22.77	0
2555	-2.449294e-16	1.000000	29.21	33.28	0.17	35.52	0

Gambar 4. Dataset Hasil Transformasi

konfigurasi model yang optimal. Pencarian parameter pada tahap pertama dilakukan menggunakan metode *Randomized Search Cross-Validation*. Validasi model menggunakan *Time Series Cross Validation* (TSCV) dengan skema $n_splits = 10$, tujuannya untuk mempertahankan karakteristik *time series* pada data. *Randomized Search* sering kali lebih efisien dalam mengeksplorasi ruang parameter yang kompleks (Filiz *et al.*, 2025). Model XGBoost memiliki perbedaan tuning dengan algoritma Random Forest. Pada XGBoost tuning dilakukan lebih intensif pada parameter *learning_rate* (*eta*) tujuannya untuk mengatur kecepatan konvergensi model. Optimasi *learning_rate* untuk meningkatkan konvergensi dan mengurangi error pada *dataset time series*. Hasil konfigurasi *hyperparameter* disajikan dalam tabel 3.

Tabel 3. Hasil *Final Tuning* Random Forest dan XGBoost

Model	Parameter	Nilai Optimal
Random Forest	<i>n_estimator</i> , <i>min_samples_split</i> , <i>min_samples_leaf</i> , <i>max_feature</i> , <i>max_depth</i> , <i>bootstrap</i>	250, 4, 5, Log2, 9, True
XGBoost	<i>n_estimator</i> , <i>learning_rate</i> , <i>colsample_bytree</i> , <i>eval_matrix</i> , <i>sub_sample</i>	250, 0.05, 0.8, logloss, 0.7



Gambar 5. Boxplot

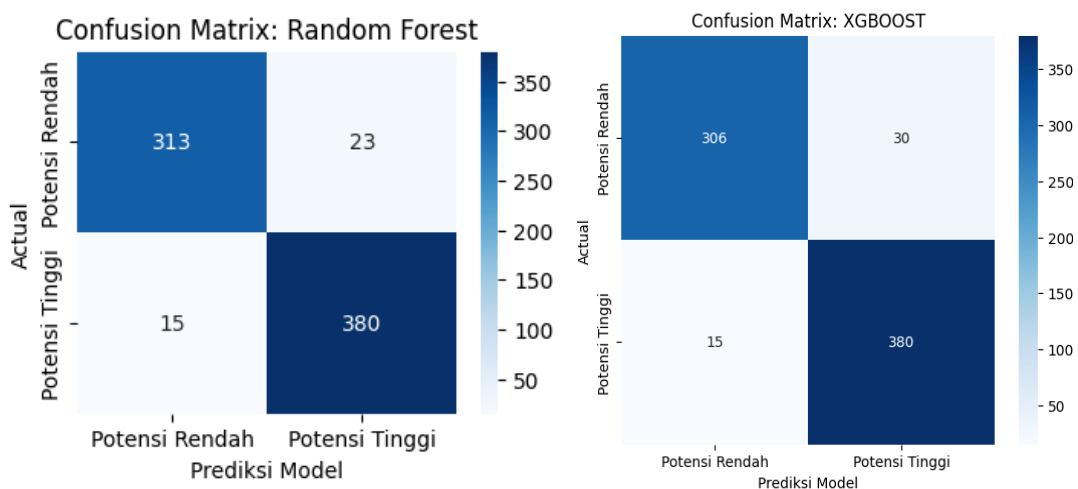
Hasil evaluasi *confusion matrix* menggunakan data pengujian untuk Random Forest, model berhasil memprediksi 313 kelas sebagai potensi rendah (0) dan 380 kelas sebagai potensi tinggi (1). Nilai *false negatif* termasuk rendah, yaitu 15 data, serta *true negatif* sebanyak 23 data. Sementara untuk XGBoost, model mampu memprediksi kelas potensi rendah (0) sebanyak 306, kelas potensi tinggi (1) sebanyak 380, *false negatif* (FN) 15, dan *true negatif* sebanyak 30. Dari *confusion matrix* tersebut menunjukkan bahwa model Random Forest memberikan hasil yang optimal dalam memprediksi kelas potensi keberadaan ikan kembung.

Menurut Lubis *et al.* (2024), nilai evaluasi confusion matrix berfungsi untuk mengetahui nilai dari matriks evaluasi, yaitu accuracy, precision, recall, dan F1-score. Berdasarkan hasil evaluasi, model Random Forest menunjukkan performa lebih unggul dibandingkan XGBoost dengan akurasi model mencapai 95%, khususnya pada identifikasi potensi tinggi (kelas 1). Dengan hasil evaluasi tersebut, Random Forest menjadi model terbaik dalam klasifikasi keberadaan ikan kembung. Secara keseluruhan, Random Forest memprediksi potensi tinggi lebih presisi, dilihat dari nilai precision (94%). Nilai F1-Score yang lebih unggul juga mengindikasikan bahwa Random Forest lebih seimbang secara keseluruhan. Meskipun kedua model menghasilkan nilai recall yang sama besar (96%), berdasarkan hasil confusion matrix, Random Forest lebih presisi dibandingkan dengan XGBoost yang lebih banyak menghasilkan false positif.

Hasil nilai matriks sejalan dengan penelitian yang dilakukan oleh Lim *et al.* (2026) tentang klasifikasi kualitas air untuk budidaya perikanan, yang menunjukkan bahwa Random Forest mencapai akurasi klasifikasi paling optimal dibandingkan dengan algoritma XGBoost dan Naive Bayes. Penelitian lain dalam klasifikasi kualitas air juga menunjukkan bahwa Random Forest lebih unggul dibandingkan dengan 7 algoritma *machine learning* lainnya termasuk XGBoost (Elmeftahi *et al.*, 2024). Random Forest cenderung lebih stabil dibandingkan dengan XGBoost untuk sistem peringatan dini dalam klasifikasi kualitas air pada budidaya ikan patin (Hadyan *et al.*, 2026). Meskipun konteks aplikasinya berbeda, keunggulan stabilitas Random Forest secara algoritmik konsisten ditemukan dalam berbagai penelitian klasifikasi berbasis data lingkungan.

Tabel 4. Evaluasi Nilai Matrix

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	95%	94%	96%	95%
XGBoost	94%	93%	96%	94%



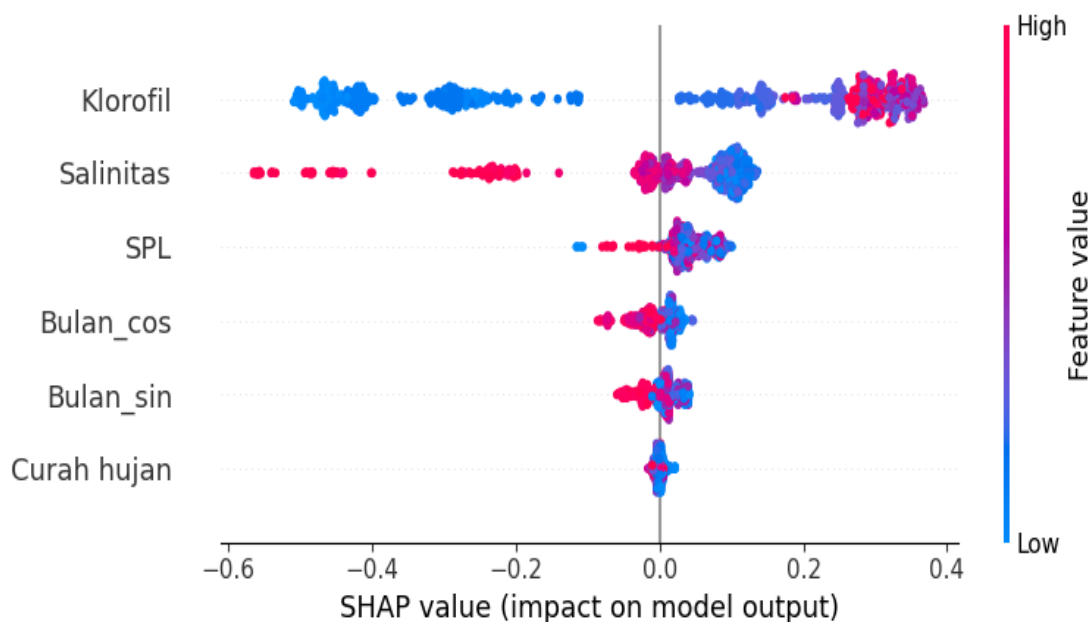
Gambar 6. Confusion Matrix Random Forest dan XGBoost

Keunggulan Random Forest dalam klasifikasi potensi keberadaan ikan kembung dapat dikaitkan dengan karakteristik dataset oseanografi. Data oseanografi harian di Selat Sunda bersifat fluktuatif (gambar 3) akibat dinamika kondisi laut yang dipengaruhi musim dan pola angin. Mekanisme *bagging* pada Random Forest terbukti lebih *robust* terhadap data yang fluktuasi dibandingkan mekanisme *boosting* pada XGBoost yang lebih sensitif terhadap *noise* pada dataset fluktuatif karena sifatnya yang membangun pohon secara sekuensial (Carreira-Perpiñán & Zharmagambetov, 2020)

Analisis SHAP

Menurut Gani *et al.* (2024) analisis SHAP bertujuan untuk melihat perilaku variabel *input* dalam menentukan hasil prediksi model. Grafik SHAP pada gambar 7 hasil analisis menggunakan model terbaik Random Forest menunjukkan bahwa klorofil-a merupakan variabel paling dominan dalam menentukan prediksi keberadaan ikan kembung. Titik merah yang terkonsentrasi ke arah positif dan titik biru yang terkonsentrasi ke sisi negatif mengindikasikan bahwa semakin tinggi nilai klorofil-a berkontribusi dalam mendorong prediksi ke arah kelas 1 (potensi tinggi). Klorofil-a merupakan indikator paling penting dalam produktivitas primer, dan dari indikator inilah kondisi ekosistem suatu perairan dapat dilihat, termasuk potensi keberadaan ikan. Klorofil-a adalah pigmen fitoplankton yang merupakan sumber makanan ikan (Sitorus *et al.*, 2022).

Salinitas di urutan ke dua, sebaran titik biru yang didominasi sisi positif mengindikasikan bahwa salinitas menuju rendah cenderung meningkatkan peluang potensi tinggi untuk keberadaan ikan kembung, sedangkan sebaran titik merah yang didominasi di sisi negatif menunjukkan bahwa semakin tinggi salinitas akan menurunkan potensi keberadaan ikan. Salinitas merupakan kadar garam terlarut dalam perairan. Parameter salinitas berperan penting dalam menentukan sebaran ikan pelagis kecil khususnya ikan kembung (Apriansyah & Utami, 2024). Selat Sunda merupakan perairan pertemuan massa air Laut Jawa yang bersalinitas rendah dan massa air Samudra Hindia yang bersalinitas tinggi. Pada Musim timur pengaruh massa air yang terbawa dari Laut Jawa menuju Samudra Hindia lebih kuat, menyebabkan perairan Selat Sunda didominasi suhu hangat dan bersalinitas rendah yang mendukung agregasi ikan kembung (Aji *et al.*, 2022). Lingkungan dengan salinitas rendah, konsumsi energi untuk osmoregulasi menurun yang berpotensi meningkatkan laju pertumbuhan pada spesies ikan kembung dan kerapu malabar (Zhu *et al.*, 2023).

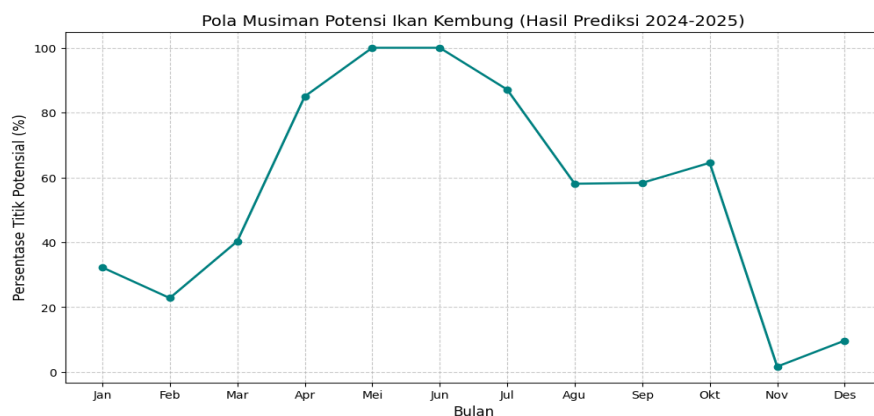


Gambar 7. SHAP Value

Variabel SPL menunjukkan pengaruh *nonlinear* terhadap keberadaan ikan kembung. Nilai SPL yang hangat berkontribusi positif terhadap prediksi potensi tinggi, namun terdapat sebaran nilai SPL tinggi yang memberikan pengaruh negatif. Sebaran titik biru di kedua sisi juga mengindikasikan bahwa penurunan suhu pada rentang tertentu justru berkontribusi positif terhadap kemunculan ikan kembung. Hal ini menunjukkan bahwa habitat ikan kembung di Selat Sunda berada pada rentang suhu yang moderat dan didukung oleh ketersediaan pakan bukan sebatas korelasi *linear* positif. Ikan kembung bermigrasi ke perairan dengan suhu yang lebih hangat serta kaya nutrisi, mengindikasikan suhu hangat moderat mendukung keberadaan ikan kembung (Hanafi *et al.*, 2025). Hubungan penurunan suhu dengan kenaikan klorofil-a sebagai parameter yang optimum untuk potensi keberadaan ikan kembung mempresentasikan fenomena *upwelling*. Fenomena *upwelling* membawa massa air kaya nutrisi dari lapisan bawah ke permukaan Selat Sunda (Ramadlanie *et al.*, 2022). Variabel bulan (sin/cos) dan curah hujan memiliki kontribusi tambahan dalam menangkap pola musiman. Curah Hujan memiliki pengaruh paling minimal, mengindikasikan bahwa dinamika stok ikan kembung di Selat Sunda lebih dipengaruhi oleh ketersediaan pakan dan stabilitas massa air.

Pola Musiman Potensi Keberadaan Ikan Kembung

Gambar 8, merupakan hasil pengujian menggunakan model terbaik Random Forest untuk prediksi data tahun 2024 - 2025. Pada grafik tersebut dapat dilihat bahwa musim pola keberadaan ikan kembung mulai terlihat di bulan April dan berada dipuncak keberadaan pada bulan Mei hingga Juni. Kondisi tersebut mengindikasikan bahwa pada bulan Mei hingga Juni parameter lingkungan seperti klorofil-a, salinitas, SPL serta intensitas curah hujan di perairan Selat Sunda berada pada nilai optimum untuk mendukung keberadaan ikan kembung. Hal tersebut sejalan dengan penelitian yang dilakukan oleh Yonviter *et al.* (2020) bahwa musim penangkapan ikan kembung di Selat Sunda terjadi pada bulan April hingga Juni. Jika dihubungkan dengan grafik variabilitas oseanografi di Selat Sunda pada gambar 3 terlihat bahwa klorofil-a pada bulan April hingga Juni memiliki rata-rata 0.2 - 0.4 mg/m³, suhu di rata-rata 29.3 - 31.0°C, salinitas 31.52 - 32.87 psu, dan rata-rata curah hujan yang relatif rendah mengindikasikan bahwa pada puncak potensi keberadaan ikan kembung, kondisi perairan berada pada rentang nilai optimum untuk mendukung keberadaan ikan kembung di Selat Sunda. Menurut hasil observasi bersama nelayan ikan kembung yang melakukan *trip* hingga ke perairan Selat Sunda, musim penangkapan ikan kembung terjadi pada musim timur atau bulan Juni. Hasil lapangan tersebut memvalidasi pola musiman yang dihasilkan model Random Forest. Transisi monsun pertama (April-Mei) menciptakan kondisi lingkungan yang optimal, seperti peningkatan suhu permukaan laut dan konsentrasi klorofil-a, yang mendukung keberadaan ikan pelagis kecil (Jha *et al.*, 2024).



Gambar 8. Grafik Pola Musiman Potensi Keberadaan Ikan Kembung

Pada bulan Juli hingga Agustus dan September terlihat grafik potensi keberadaan ikan kembung mulai menurun, hal ini disebabkan oleh kondisi perairan yang berubah seperti pola arus, pola angin dan cuaca. Periode musim barat yang terjadi pada bulan November hingga Maret, angin bertiup dari barat laut menuju tenggara mengakibatkan terjadinya peningkatan arus di permukaan sehingga meningkatkan risiko bagi aktivitas pelayaran (Daud *et al.*, 2025). Perubahan arah dan kecepatan arus laut juga berdampak terhadap konsentrasi sebaran ikan terutama spesies kembung dan tongkol (Marlina, 2025). Meskipun pada gambar 3 terlihat bahwa pola musiman klorofil-a terjadi pada bulan Desember pada nilai tertingginya, namun grafik pola musiman potensi keberadaan ikan kembung justru berada pada musim paceklik atau musim penurunan keberadaan, ini terjadi karena meskipun klorofil-a merupakan variabel utama di analisis SHAP dalam menentukan prediksi keberadaan ikan kembung, kondisi variabel lain seperti suhu yang menurun serta fluktuasi salinitas di musim barat yang menyebabkan kondisi perairan tidak stabil sehingga menjadi kurang ideal untuk kehidupan ikan kembung meskipun makanan melimpah. Nilai klorofil-a yang terlalu tinggi justru mengakibatkan kekeruhan di perairan, curah hujan yang tinggi mendorong masuknya nutrisi dari daratan yang juga membawa sedimentasi (Graha *et al.*, 2024). Kondisi tersebut berpotensi memengaruhi aktivitas mencari makan ikan kembung, mengingat ikan famili *Scombridae* termasuk ikan kembung umumnya mengandalkan penglihatan untuk mencari makan (Pratiwi *et al.*, 2025).

Rendahnya potensi keberadaan ikan kembung di bulan November dan Desember juga dipengaruhi oleh hubungan fisiologi langsung antara ikan kembung dan salinitas. Jika dilihat pada gambar 3, pada bulan November menunjukkan bahwa pola konsentrasi salinitas cenderung tinggi. Hal tersebut juga mengonfirmasi hasil dari analisis SHAP, kondisi salinitas tinggi cenderung memberikan kontribusi negatif terhadap keberadaan ikan kembung. Menurut Pratama *et al.* (2023) ikan pelagis kecil sangat rentan terhadap perubahan salinitas, sehingga perubahan salinitas diluar batas optimum akan mengurangi kesesuaian habitat. Secara keseluruhan, model berhasil mengidentifikasi periode utama musim keberadaan ikan kembung yang juga berkorelasi dengan peningkatan tangkapan ikan kembung di perairan Selat Sunda terjadi pada bulan April hingga Juni, sedangkan untuk periode paceklik atau penurunan keberadaan ikan kembung terjadi pada bulan November hingga Februari.

KESIMPULAN

Hasil perbandingan evaluasi dari model Random Forest dan XGBoost menggunakan pendekatan time series untuk klasifikasi potensi keberadaan ikan kembung menunjukkan algoritma Random Forest lebih unggul dalam memodelkan klasifikasi potensi keberadaan ikan kembung di Selat Sunda. Dengan akurasi 95%, model Random Forest dapat mengklasifikasikan kelas potensi tinggi dengan nilai recall sebesar 96% yang mengindikasikan bahwa model Random Forest mampu mengenali pola habitat potensi tinggi dari total data pengujian. Sementara itu model XGBoost juga menghasilkan akurasi model yang optimal di angka 94% dengan nilai *recall* 96%. Namun, diperlukan proses tuning yang parameter yang lebih kompleks untuk mencapai performa optimal. Analisis SHAP menunjukkan klorofil-a sebagai variabel yang paling dominan dalam menentukan prediksi keberadaan ikan kembung dengan korelasi hubungan yang positif, namun pengaruhnya terhadap keberadaan ikan kembung dipengaruhi oleh faktor lain seperti salinitas di urutan kedua kemudian SPL. Pola musiman potensi keberadaan ikan kembung menunjukkan bahwa puncak keberadaan ikan kembung di Selat Sunda terjadi pada bulan Mei hingga Juni dan mengalami paceklik di bulan November hingga Februari. Namun, penelitian ini memiliki keterbatasan yaitu dominasi klorofil-a dalam analisis SHAP kemungkinan dipengaruhi oleh skema pelabelan yang menggunakan klorofil-a sebagai kriteria utama, sehingga interpretasi SHAP perlu mempertimbangkan hal tersebut.

DAFTAR PUSTAKA

Agustina, F.D., Arif, M., & Ahmad, S., 2025. Systematic literature review atas kinerja algoritma KNN, Naïve Bayes, dan Decision Tree pada berbagai studi prediksi dan klasifikasi. *Jurnal Jawara Sistem Informasi*, 3(1):1–9.

- Aji, T., Pranowo, W.S., Santosa, Y.N., Hendra, H., & Umam, C., 2022. Analisis massa air musiman di Selat Sunda: Seasonal water mass analysis in the Sunda Strait. *Jurnal Chart Datum*, 8(2):125–142. DOI: 10.37875/chartdatum.v8i2.148
- Akita, E.A., Gaol, J.L., & Amri, K., 2022. Model maximum entropy untuk prediksi daerah penangkapan ikan pelagis kecil di Laut Jawa. *Jurnal Ilmu dan Teknologi Kelautan Tropis*, 14(3):449–461. DOI: 10.29244/jitkt.v14i3.45164
- Apriansyah, & Utami, P.B., 2024. Karakteristik oseanografi dan musim penangkapan ikan kembung (*Rastrelliger* sp.) di Laut Natuna. *Manfish Jurnal*, 5(1):22–31.
- Asbar, & Ihsan, 2022. Pemetaan daerah penangkapan ikan pelagis kecil untuk meningkatkan hasil tangkapan nelayan di perairan Kota Makassar. *Jurnal Pengabdian Masyarakat Kauniah*, 1(1):1–13. DOI: 10.33096/jamka.v1i1.110
- Baldé, B.S., Döring, J., Ekau, W., Diouf, M., & Brehmer, P., 2019. Bonga shad (*Ethmalosa fimbriata*) spawning tactics in an upwelling environment. *Fisheries Oceanography*, 28(6):686–697. DOI: 10.1111/fog.12451
- Caesario, R., Delis, P.C., & Julian, D., 2022. Struktur ukuran, tipe pertumbuhan dan faktor kondisi ikan kembung lelaki (*Rastrelliger kanagurta*) yang didaratkan di Pelabuhan Perikanan Pantai Lempasing. *Jurnal Akuatika Indonesia*, 7(2):87–92.
- Carreira-Perpiñán, M.Á., & Zharmagambetov, A., 2020. Ensembles of bagged TAO trees consistently improve over random forests, AdaBoost and gradient boosting. *Proceedings of the 2020 ACM-IMS Foundations of Data Science Conference*, 35–46. DOI: 10.1145/3412815.3416882
- Daud, A., Pratomo, D.G., Khomsin, & Putra, A.W., 2025. Identifikasi faktor-faktor yang memengaruhi kecepatan arus laut di perairan Selat Sunda. *Lambda: Jurnal Ilmiah Pendidikan MIPA dan Aplikasinya*, 5(2):386–394. DOI: 10.58218/lambda.v5i2.1366
- Dawolo, A.J., Lase, R.C., Waruwu, J.C., Gea, R., Harefa, S.V., & Zai, S., 2025. Analisis pertumbuhan dan tingkat kematangan gonad ikan kembung (*Rastrelliger*) untuk mendukung pengelolaan perikanan berkelanjutan. *Zoologi: Jurnal Ilmu Peternakan, Ilmu Perikanan, Ilmu Kedokteran Hewan*, 3(1):89–100. DOI: 10.62951/zoologi.v3i1.110
- Dunstan, P.K., Foster, S.D., King, E., Risbey, J., O’Kane, T.J., Monselesan, D., Hobday, A.J., Hartog, J.R., & Thompson, P.A., 2018. Global patterns of change and variation in sea surface temperature and chlorophyll-a. *Scientific Reports*, 8(1):1-9. DOI: 10.1038/s41598-018-33057
- Durap, A., 2025. Predicting ocean parameters with explainable machine learning: Overcoming scale and time challenges. *Regional Studies in Marine Science*, 90:1–32.
- Dwiyanti, A., Maslukah, L., & Rifai, A., 2022. Pengaruh suhu permukaan laut (SPL) dan klorofil-a terhadap hasil tangkapan ikan layang (*Decapterus macrosoma*) di perairan Kabupaten Rembang, Jawa Tengah. *Indonesian Journal of Oceanography*, 4:109–119. DOI: 10.14710/ijoce.v4i4.15708
- Elmeftahi, S., Rakhman, M.D., & Rahmatulloh, A., 2024. Comparison of machine learning algorithms in detecting contaminants in drinkable water. *Innovation in Research of Informatics (INNOVATICS)*, 6(1):7–14.
- Ernawati, A., Sitorus, Z., Iqbal, M., & Nasution, D., 2025. Penerapan data mining untuk klasifikasi penduduk miskin di Kabupaten Labuhanbatu menggunakan Random Forest dan K-Nearest Neighbors. *Bulletin of Information Technology*, 6(2):23–35. DOI: 10.47065/bit.v5i2.1783
- Fadli, M., & Saputra, R.A., 2023. Klasifikasi dan evaluasi performa model Random Forest untuk prediksi stroke. *JT: Jurnal Teknik*, 12(2):72–80.
- Fauzi, N.P., Khomsah, S., & Wicaksono, A.D.P., 2025. Penerapan feature engineering dan hyperparameter tuning untuk meningkatkan akurasi model Random Forest pada klasifikasi risiko kredit. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 12(2):251–262. DOI: 10.25126/jtiik.2025128472
- Filiz, G., Kara, E., Altıntaş, S., Drias, Y., Yıldız, A., & Çakar, T., 2025. Comparison of deep learning and tree-based models for internet traffic speed prediction. *2025 33rd Signal Processing and Communications Applications Conference*, 1–4. DOI: 10.1109/SIU66497.2025.11112468
- Firdaus, M. et al., 2018. The profile of tuna and cakalang fishery in Indonesia. *Buletin Ilmiah Marina Sosial Ekonomi Kelautan dan Perikanan*, 4(1):23–31.

- Gani, H., Damayanti, A.D., Mubarak, M.M., Gani, H., & Muslimin, 2024. Model pembelajaran mesin yang dapat dijelaskan untuk menjelaskan parameter iklim yang berpengaruh berdasarkan prediksi curah hujan. *Jurnal IT: Media Informasi IT STMIK Handayani*, 15(2):98–110. DOI: 10.37639/jti.v15i2.378
- Graha, D.S., Kunarso, & Zainuri, M., 2024. Variabilitas antar tahunan suhu permukaan laut dan klorofil-a terhadap hasil tangkapan ikan kembung (*Rastrelliger* sp.) di perairan Demak. *Indonesian Journal of Oceanography*, 6(2):148–158. DOI: 10.14710/ijoce.v6i2.17895
- Hadyan, M.R., Marleny, F.D., & Ningrum, A.A., 2026. Enhancing water quality early warning system accuracy in *Pangasius* aquaculture using machine learning. *Journal of Computer Networks, Architecture and High Performance Computing*, 8(1):10–19. DOI: 10.47709/cnahpc.v8i1.7479
- Hanafi, H., Aprinsyah, A., & Safitri, I., 2025. Variasi musiman suhu dan klorofil-a serta hubungan dengan hasil tangkapan ikan kembung (*Rastrelliger* sp.) Laut Natuna. *Jurnal Kelautan Tropis*, 28(1):164–174. DOI: 10.14710/jkt.v28i1.25600
- Ilaihiyah, R., Abdurrahman, G. and Daryanto., 2026. Implementasi Algoritma XGBoost Classifier untuk Klasifikasi Kerumunan Jemaah Haji dan Umrah. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 6(1):106–115. DOI:10.57152/malcom.v6i1.2245
- Jha, S., Sudhakar, S.D., Majumder, S., Joseph, S., & Nair, T.M.B., 2024. A multidecadal study of the Malabar upwelling system influencing Indian mackerel landings along the coasts of Karnataka and Kerala, south-east Arabian Sea. *Journal of Marine Systems*, 246. DOI: 10.1016/j.jmarsys.2024.104012
- Kurnianto, A., Sitanggang, I.S., Kusuma, M., & Hardhienata, D., 2024. Klasifikasi daerah penangkapan ikan menggunakan algoritma Random Forest dan Support Vector Machine. *Jurnal Ilmu Komputer Agri-Informatika*, 11(2):100–110. DOI: 10.29244/jika.11.2.100-110
- Lim, G.J., Putra, J.C., Agusia, P., & Utama, M.Y., 2026. Perbandingan Naive Bayes, Random Forest, XGBoost untuk klasifikasi mutu air. *Jurnal Ilmu Komputer dan Sistem Informasi*, 14(1):1–9. DOI: 10.24912/03ch2p08
- Lubis, M.G.A., Palupi, R., Astriani, A., & Khotimah, P.A., 2024. Model prediksi kelimpahan ikan tongkol (*Euthynnus affinis*) menggunakan metode K-Nearest Neighbor (K-NN) di Laut Jawa. *Journal of Fisheries and Marine Research*, 8(3):1–7.
- Luangmaneerote, S., & Thasana, W., 2025. Boosting medium-size reservoir water level predictions using cyclical encoding. *Engineering, Technology and Applied Science Research*, 15(3):23672–23679. DOI: 10.48084/etasr.10595
- Marlina, S., 2025. Analisis pengaruh perubahan pola arus laut terhadap distribusi ikan tangkap. *Jurnal Perikanan Indonesia*, 1(1):29–35.
- Prasetyo, T., Putri, R.A., Ramadhani, D., Angraini, Y., & Notodiputro, K.A., 2024. Perbandingan kinerja metode ARIMA, Multi-Layer Perceptron, dan Random Forest dalam peramalan harga logam mulia berjangka yang mengandung pencilan. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(2):265–274. DOI: 10.25126/jtiik.20241127392
- Pratiwi, A., Almira, C.S., Nasir, F., Abidin, M., & Muchlisin, Z.A., 2025. Literature review: Komposisi lambung famili *Scombridae* di perairan Indonesia. *Zoologi: Jurnal Ilmu Peternakan, Ilmu Perikanan, Ilmu Kedokteran Hewan*, 3(1):19–28. DOI: 10.62951/zoologi.v3i1.92
- Pratama, G.B., Nurani, T.W., Mustaruddin, & Herdiyeni, Y., 2022. Modeling habitat suitability of pelagic fish based on oceanographic conditions in Palabuhanratu waters. *Bawal Widya Riset Perikanan Tangkap*, 14(3):161–171. DOI: 10.15578/bawal.14.3.2022.161-171
- Purba, R., Handoco, E., & Sinaga, M.P., 2025. Pengaruh suhu permukaan laut (SPL) dan kandungan klorofil-a terhadap hasil tangkapan ikan kembung (*Rastrelliger kanagurta*) yang didaratkan di Pelabuhan Perikanan Samudra (PPS) Belawan Provinsi Sumatera Utara. *Jurnal Penelitian Teknik dan Sumber Daya Perairan*, 1(1):28–35.
- Purnawan, S., Manik, H.M., Hestirianoto, T., Ulfah, M., & Amri, K., 2025. Demersal fish distribution based on the integration of seabed acoustic and environmental features using machine learning in Northern Aceh, Indonesia. *Regional Studies in Marine Science*, 91:1–13. DOI: 10.1016/j.rsma.2025.104563

- Puspasari, R., Triharyuni, S., Alimi, T., Campbell, S.J., Jakub, R., Suherfian, W., de la Rosa, E. & Setiawan, H., 2021. Pengaruh ENSO terhadap lingkungan perairan dan perikanan di perairan Sulawesi Utara, *Jurnal Penelitian Perikanan Indonesia*, 27(2):95–106. DOI:10.15578/jppi.27.2.2021.95 -106
- Ramadlanie, F.Y., Indrayanti, E., & Handoyo, G., 2022. Identifikasi daerah upwelling berdasarkan indikator suhu permukaan laut dan klorofil-a di perairan Selat Sunda. *Indonesian Journal of Oceanography (IJOCE)*, 4(4):62–73. DOI: 10.14710/ijoce.v4i4.15341
- Rashid, L.B., Shuja, S.Z., & Rehman, S., 2025. Machine learning forecasting of direct solar radiation: A multi-model evaluation with trigonometric cyclical encoding. *Forecasting*, 7(4):58–82. DOI: 10.3390/forecast7040058
- Rizky, M., & Andarsyah, R., 2023. Klasifikasi MIT-BIH Arrhythmia database metode Random Forest dan CNN dengan model ResNet-50: A systematic literature review. *Jurnal Teknologi dan Sistem Informasi Bisnis*, 5(3):190–196. DOI: 10.47233/jteksis.v5i3.825
- Sitorus, J.H., Atmojo, A.T., Bachri, S., Prayitno, H.S., & Komarita, I., 2022. Analysis of fishing potential zones based on SST, chlorophyll-a, and boat detection and assessing RZWP3K, Lampung. *Jurnal Teknologi Perikanan dan Kelautan*, 13(1):89–102. DOI: 10.24319/jtpk.13.89-102
- Sihombing, S.F., Pakpahan, J.P. & Lubis, A.H., 2025. Klasifikasi Produk Iphone dengan Menggunakan Algoritma XGBoost, *Journal of Informatics Management and Information Technology*, 5(3):371–380. DOI: 10.47065/jimat.v5i3.649
- Tarigan, D.J., Cahyadi, F.D., Sasongko, A.S., Yonanto, L., & Rahayu, B.D., 2020. Daerah penangkapan ikan kembung (*Rastrelliger* sp.) di Selat Sunda pada musim peralihan. *Jurnal Teknologi Perikanan dan Kelautan*, 11(1):63–79. DOI: 10.24319/jtpk.11.63-79
- Tarigan, D.J., Saleha, A.N., & Rahayu, B.D., 2025. Potensi zona perikanan *Rastrelliger* sp. di Selat Sunda. *Jurnal Pendidikan Perikanan Kelautan*, 3(2):60–66.
- Valerian, F.R., Syarief, M., & Fatah, D.A., 2025. Klasifikasi tingkat obesitas menggunakan metode GBM dan confusion matrix. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2):2242–2249.
- Yulianti, S.E.H., Soesanto, O., & Sukmawaty, Y., 2022. Penerapan metode Extreme Gradient Boosting (XGBOOST) pada klasifikasi nasabah kartu kredit. *Journal of Mathematics: Theory and Applications (JOMTA)*, 4(1):21–26.
- Zhu, Y., Negishi, R., Fukunaga, K., Udagawa, S., Shimabukuro, A., et al., 2023. Activation of the growth-IGF-1 axis, but not appetite, is related to high growth performance in juveniles of the Malabar grouper, *Epinephelus malabaricus*, under isosmotic condition. *Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology*. DOI: 10.1016/j.cbpa.2023.111456
- Zulfa, I.N., Wirasatriya, A., & Ismanto, A., 2024. Kajian spasial dan temporal klorofil-a di Selat Makassar: Variasi musiman dan antar tahunan. *Buletin Oseanografi Marina*, 13:311–326. DOI: 10.14710/buloma.v13i3.63067