

**PENGEMBANGAN MODEL DETEKSI DINI POTENSI TUMPAHAN MINYAK
MENGUNAKAN ALGORITMA XGBOOST BERDASARKAN DATA SAR**
**Development of an Early Detection Model for Potential Oil Spills Using the XGBoost
Algorithm Based on SAR Data**

Najla Syafa Kamila^{1*}, Alzechtra Deffanola² dan Jihad Izdihar Khoiron³

¹²³Marine Information System, Kampus Serang, Universitas Pendidikan Indonesia, Serang
42116, Indonesia

*Corresponding author, e-mail: najlasyafa@upi.edu

ABSTRACT

Oil spill is a form of marine pollution that can have serious impacts on coastal ecosystems and community economic activities. Synthetic Aperture Radar (SAR) technology is widely used to detect oil spills because it can operate both day and night and is not affected by weather conditions. This research aims to develop an early detection model for potential oil spills using the Extreme Gradient Boosting (XGBoost) algorithm based on features extracted from SAR images. The dataset used consists of numerical data from SAR image extraction with binary labels for oil spill and non-oil spill. To address the issue of class imbalance, the research applies the Synthetic Minority Oversampling Technique (SMOTE) method to the training data. The model's performance is evaluated using the metrics of accuracy, precision, recall, F1-score, and ROC-AUC. The research results show that the XGBoost model is capable of achieving an accuracy of 96.28%, precision of 57.14%, recall of 50.00%, F1-score of 53.33%, and ROC-AUC of 94.51%. In addition to producing binary classifications, the model is also able to generate probability values categorized into Low, Moderate, High, and Very High risk levels as a basis for monitoring priorities. The research results indicate that the combination of XGBoost and SMOTE has the potential to support the development of a faster, more automated, and efficient early oil spill detection system based on SAR data.

Keywords: Early Warning System, Oil Spill, SAR, SMOTE, XGBoost

ABSTRAK

Tumpahan minyak (*oil spill*) merupakan salah satu bentuk pencemaran laut yang dapat menimbulkan dampak serius terhadap ekosistem pesisir dan aktivitas ekonomi masyarakat. Teknologi *Synthetic Aperture Radar* (SAR) banyak digunakan untuk mendeteksi tumpahan minyak karena mampu beroperasi pada siang maupun malam hari serta tidak dipengaruhi oleh kondisi cuaca. Penelitian ini bertujuan mengembangkan model deteksi dini potensi tumpahan minyak menggunakan algoritma Extreme Gradient Boosting (XGBoost) berdasarkan fitur hasil ekstraksi citra SAR. Dataset yang digunakan terdiri atas data numerik hasil ekstraksi citra SAR dengan label biner *oil spill* dan *non-oil spill*. Untuk mengatasi permasalahan ketidakseimbangan kelas, penelitian menerapkan metode *Synthetic Minority Oversampling Technique* (SMOTE) pada data pelatihan. Performa model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan ROC-AUC. Hasil penelitian menunjukkan bahwa model XGBoost mampu mencapai nilai *accuracy* sebesar 96,28%, *precision* 57,14%, *recall* 50,00%, *F1-score* 53,33%, dan ROC-AUC 94,51%. Selain menghasilkan klasifikasi biner,

model juga mampu menghasilkan nilai probabilitas yang dikategorikan ke dalam tingkat risiko Low, Moderate, High, dan Very High sebagai dasar prioritas pemantauan. Hasil penelitian menunjukkan bahwa kombinasi XGBoost dan SMOTE berpotensi mendukung pengembangan sistem deteksi dini tumpahan minyak yang lebih cepat, otomatis, dan efisien berbasis data SAR.

Kata kunci: Sistem Deteksi Dini, Tumpahan Minyak, SAR, SMOTE, XGBoost

PENDAHULUAN

Tumpahan minyak di laut merupakan salah satu bentuk pencemaran yang memberikan dampak signifikan terhadap lingkungan pesisir dan aktivitas ekonomi masyarakat. Sumber pencemaran ini dapat berasal dari berbagai kegiatan, seperti kecelakaan kapal tanker, kebocoran pipa minyak, maupun aktivitas industri lepas pantai. Keberadaan minyak di perairan berpotensi merusak habitat penting seperti terumbu karang, mangrove, dan padang lamun, mengancam kelangsungan biota laut, serta mengganggu sektor perikanan dan pariwisata yang menjadi sumber mata pencaharian masyarakat pesisir (Supriyanto *et al.* 2025). Berbagai insiden tumpahan minyak yang terjadi di wilayah perairan strategis Indonesia, termasuk Selat Makassar dan Laut Jawa, mengindikasikan bahwa ancaman pencemaran minyak masih berlangsung dan berpotensi terulang, sehingga diperlukan mekanisme pemantauan yang andal dan berkelanjutan (Prastyani dan Basith, 2019). Mengingat kejadian tumpahan minyak sering kali tidak dapat diprediksi secara visual pada tahap awal, diperlukan sistem deteksi yang cepat, otomatis, dan dapat diandalkan untuk mendukung respons mitigasi secara dini.

Teknologi *Synthetic Aperture Radar* (SAR) telah terbukti menjadi salah satu instrumen penginderaan jauh yang paling efektif untuk mendeteksi tumpahan minyak di permukaan laut. Citra SAR banyak dimanfaatkan dalam pemantauan tumpahan minyak karena mampu mengakuisisi data pada siang dan malam hari serta tetap efektif digunakan pada berbagai kondisi cuaca, sehingga mendukung pengawasan wilayah laut secara kontinu dan luas (Suniada, 2021). Lapisan minyak pada permukaan laut meredam gelombang kapiler sehingga menghasilkan area gelap (*dark formation*) yang dapat dibedakan dari area laut normal pada citra SAR. Karakteristik inilah yang menjadi dasar ilmiah penggunaan data SAR sebagai sumber fitur utama dalam model deteksi tumpahan minyak berbasis *machine learning* (Topouzelis, 2008).

Dalam beberapa tahun terakhir, pendekatan *machine learning* telah banyak diterapkan untuk mengotomatiskan proses deteksi tumpahan minyak pada citra SAR. Bianchi *et al.* (2020) menunjukkan bahwa metode berbasis pembelajaran mesin mampu menghasilkan performa deteksi yang sebanding dengan interpretasi visual oleh manusia pada citra SAR skala besar.

Zeng dan Wang (2020) juga membuktikan bahwa algoritma kecerdasan buatan memiliki kemampuan tinggi dalam mengidentifikasi tumpahan minyak secara otomatis dan efisien. Sementara itu, Conceação *et al.* (2021) mengembangkan sistem deteksi berbasis *Random Forest* pada citra Sentinel-1 SAR dan menunjukkan bahwa pemilihan fitur yang relevan dari data SAR dapat meningkatkan kemampuan model dalam membedakan *oil spill* dari fenomena laut lainnya yang memiliki karakteristik visual serupa (*look-alikes*). Di tingkat nasional, Marsisno (2025) juga membuktikan bahwa XGBoost merupakan algoritma dengan performa terbaik untuk deteksi tumpahan minyak berbasis citra Sentinel-1 SAR di perairan Bintan, Kepulauan Riau, dengan akurasi mencapai 97,72%.

Salah satu tantangan utama dalam deteksi tumpahan minyak berbasis *machine learning* adalah ketidakseimbangan kelas (*class imbalance*) pada dataset, di mana jumlah sampel *non-oil spill* jauh lebih banyak dibandingkan sampel *oil spill*. Kondisi ini berpotensi menyebabkan model cenderung mengabaikan kelas minoritas yang justru paling penting untuk dideteksi. Sutoyo dan Fadlurranman (2020) menyatakan bahwa representasi kelas yang seimbang pada data pelatihan merupakan faktor penting yang memengaruhi kemampuan model klasifikasi dalam mengenali setiap kelas secara optimal. Ketidakseimbangan distribusi data dapat menyebabkan model lebih cenderung mempelajari karakteristik kelas mayoritas dan mengabaikan kelas minoritas. Oleh karena itu, penerapan teknik *Synthetic Minority Oversampling Technique* (SMOTE) diperlukan untuk meningkatkan representasi kelas minoritas sehingga performa klasifikasi dapat ditingkatkan.

Penelitian ini bertujuan untuk mengembangkan model deteksi dini potensi tumpahan minyak menggunakan algoritma *Extreme Gradient Boosting* (XGBoost) berdasarkan fitur hasil ekstraksi citra SAR. Dataset yang digunakan berisi fitur numerik yang diekstrak dari citra SAR beserta label biner yang menunjukkan keberadaan *oil spill* atau *non-oil spill*. Untuk mengatasi ketidakseimbangan kelas, penelitian ini menerapkan metode SMOTE pada data pelatihan. Performa model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan ROC-AUC, serta dilengkapi dengan analisis *confusion matrix* dan *feature importance*. Selain menghasilkan klasifikasi biner, model juga menghasilkan nilai probabilitas yang dikategorikan ke dalam tingkat risiko (*Low*, *Moderate*, *High*, dan *Very High*) sehingga model berpotensi dimanfaatkan sebagai komponen pendukung sistem deteksi dini (*early warning system*) tumpahan minyak berbasis data SAR. Hasil penelitian ini diharapkan berkontribusi pada pengembangan sistem pemantauan lingkungan laut yang lebih otomatis, cepat, dan efisien sebagaimana yang direkomendasikan oleh Al-Ruzouq *et al.* (2020).

METODE PENELITIAN

2.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif berbasis machine learning untuk mengembangkan model deteksi dini potensi tumpahan minyak (*oil spill*) berdasarkan data ekstraksi citra *Synthetic Aperture Radar* (SAR). Alur penelitian terdiri dari lima tahap utama, yaitu: pengumpulan data, *pre-processing*, penyeimbangan data (*data balancing*), pelatihan model klasifikasi, dan evaluasi model.

2.2.

Dataset

Dataset yang digunakan dalam penelitian ini adalah *oil_spill.csv*, yang berisi fitur-fitur numerik hasil ekstraksi citra SAR beserta satu kolom target biner yang merepresentasikan keberadaan *oil spill* (label 1) atau *non-oil spill* (label 0). Dataset dimuat menggunakan pustaka *pandas* dan diperiksa terhadap nilai hilang (*missing values*), distribusi kelas, serta karakteristik statistik awal sebelum masuk ke tahap pemrosesan lebih lanjut.

2.3.

Pre-processing

Data

Tahap *preprocessing* mencakup pemisahan fitur dan label target, serta konversi label apabila kolom target masih bertipe teks (*object*). Apabila terdapat dua nilai unik pada kolom target, nilai pertama dipetakan ke 0 dan nilai kedua dipetakan ke 1 secara otomatis menggunakan fungsi `map()`. Selanjutnya, data dibagi menjadi data latih (80%) dan data uji (20%) menggunakan fungsi `train_test_split` dari pustaka *scikit-learn*, dengan parameter `stratify=y` untuk menjaga proporsi distribusi kelas pada kedua subset data.

2.4.

Penyeimbangan

Data

dengan

SMOTE

Dataset SAR untuk deteksi tumpahan minyak umumnya memiliki ketidakseimbangan kelas (*class imbalance*), di mana jumlah sampel *non-oil spill* jauh lebih banyak dibandingkan sampel *oil spill*. Pada *dataset* yang digunakan, distribusi kelas sebelum penyeimbangan menunjukkan 716 sampel kelas 0 dan 33 sampel kelas 1. Untuk mengatasi ketidakseimbangan ini, penelitian menerapkan metode *Synthetic Minority Oversampling Technique* (SMOTE) dari pustaka *imbalanced-learn* dengan `random_state=42`. SMOTE bekerja dengan membuat sampel sintetis baru pada kelas minoritas secara interpolasi di antara sampel-sampel yang sudah ada, sehingga setelah proses penyeimbangan jumlah data pada kedua kelas menjadi setara (716 sampel tiap kelas).

2.5. Pemodelan XGBoost

Algoritma yang digunakan untuk klasifikasi adalah *Extreme Gradient Boosting* (XGBoost), yang diimplementasikan menggunakan kelas `XGBClassifier` dari pustaka `xgboost`. XGBoost dipilih karena kemampuannya membangun pohon keputusan secara bertahap (*gradient boosting*), efektivitasnya pada dataset berukuran sedang, serta ketahanannya terhadap *overfitting*. Model dikonfigurasi dengan parameter sebagai berikut: `n_estimators=300`, `max_depth=5`, `learning_rate=0.05`, `subsample=0.8`, `colsample_bytree=0.8`, `random_state=42`, dan `eval_metric='logloss'`. Model dilatih menggunakan data latih hasil SMOTE (`X_train_smote`, `y_train_smote`).

2.6. Pemodelan Machine Learning

Evaluasi model dilakukan pada data uji yang tidak pernah dilihat oleh model selama proses pelatihan. Metrik evaluasi yang digunakan mencakup *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan ROC-AUC, yang semuanya dihitung menggunakan modul `sklearn.metrics`. Selain itu, *confusion matrix* divisualisasikan menggunakan `seaborn.heatmap` dan kurva ROC ditampilkan menggunakan `RocCurveDisplay.from_predictions()`. Hasil evaluasi model disajikan pada

Tabel 1. Hasil Evaluasi Model XGBoost	
Metrik Evaluasi	Nilai
<i>Accuracy</i>	0.9628
<i>Precision (Oil Spill)</i>	0.5714
<i>Recall (Oil Spill)</i>	0.5
<i>F1-Score (Oil Spill)</i>	0.5333
ROC-AUC	0.9451

Tabel 1 dan Tabel 2.

Tabel 2. Confusion Matrix Model XGBoost pada Data Uji

	Prediksi <i>Non-Oil Spill</i> (0)	Prediksi <i>Oil Spill</i> (1)	Total
Aktual <i>Non-Oil Spill</i> (0)	177 (TN)	3 (FP)	180
Aktual <i>Oil Spill</i> (1)	4 (FN)	4 (TP)	8

2.7. Analisis *Feature Importance*

Analisis *feature importance* dilakukan untuk mengidentifikasi fitur-fitur yang paling berpengaruh dalam proses pengambilan keputusan model. Nilai kepentingan setiap fitur diekstrak dari atribut `model.feature_importances_` dan diurutkan secara menurun. Sepuluh fitur teratas divisualisasikan dalam bentuk *bar chart horizontal* menggunakan pustaka `seaborn`, yang memungkinkan identifikasi fitur dominan dalam klasifikasi *oil spill* versus *non-oil spill*.

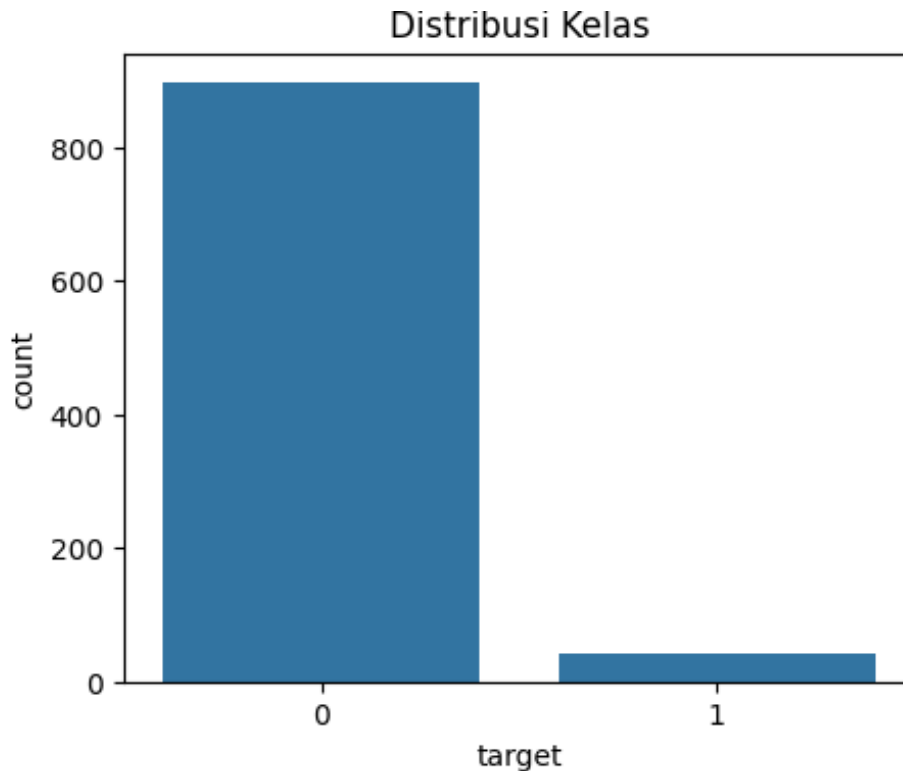
2.8. Klasifikasi Risiko Sebagai Early Warning System

Selain menghasilkan prediksi biner, model juga menghasilkan nilai probabilitas *oil spill* melalui fungsi `predict_proba()`. Nilai probabilitas tersebut kemudian dikategorikan ke dalam empat tingkat risiko menggunakan fungsi `risk_level()` dengan ambang batas (*threshold*) sebagai berikut: *Low* (probabilitas < 0.30), *Moderate* ($0.30 - 0.59$), *High* ($0.60 - 0.79$), dan *Very High* (≥ 0.80). Pendekatan ini memungkinkan model digunakan sebagai komponen early warning system yang tidak hanya memberikan keputusan biner, tetapi juga menunjukkan tingkat urgensi pemantauan berdasarkan probabilitas yang dihasilkan.

HASIL DAN PEMBAHASAN

Salah satu masalah utama dalam penelitian deteksi tumpahan minyak berbasis ML adalah ketidakseimbangan kelas. *Dataset* yang digunakan memiliki lebih banyak sampel *non-oil spill* daripada sampel *oil spill*, sehingga model mungkin lebih dominan dalam mempelajari pola dari kelas mayoritas. Untuk menyelesaikan masalah ini, penelitian ini menggunakan teknik *oversampling* minoritas sintetis (SMOTE), yang meningkatkan representasi kelas pelempar minyak pada data pelatihan. Menurut Al-Ruzouq *et al.* (2020), dua komponen penting yang memengaruhi keberhasilan algoritma pembelajaran mesin dalam mendeteksi tumpahan minyak

secara otomatis adalah kualitas data pelatihan dan representasi kelas yang memadai. Hasil penelitian menunjukkan bahwa proses *balancing* membantu model mengenali karakteristik cipratan minyak dengan lebih baik. Akibatnya, kinerja klasifikasi dapat ditingkatkan. Karakteristik dataset dapat dilihat pada Gambar 1.



Gambar 1. Karakteristik Dataset

Hasil evaluasi model dapat dilihat pada Gambar 2. Model ini menunjukkan nilai ketepatan sebesar 96,28%, yang menunjukkan bahwa model mampu membedakan data kelas 0 dan kelas 1, dan nilai ketepatan yang tinggi menunjukkan bahwa model secara keseluruhan mampu membedakan data kelas 0 dan kelas 1. Namun, karena jumlah data di kedua kelas tidak seimbang, nilai ketepatan saja belum cukup untuk menunjukkan kemampuan model secara keseluruhan. Oleh karena itu, analisis harus dilakukan menggunakan metrik tambahan seperti ketepatan, *recall*, dan skor F1.

Hasil pengujian menunjukkan nilai *precision* kelas 1 sebesar 57,14%, yang berarti sekitar 57% dari seluruh data kelas 1 yang diprediksi benar. Sementara itu, nilai *recall* sebesar 50% menunjukkan bahwa model hanya dapat menemukan setengah dari total data kelas 1 yang sebenarnya. Dengan nilai F1-score sebesar 53,33%, *precision* dan *recall* seimbang dalam mendeteksi kelas minoritas. Selain itu, nilai rata-rata *precision* sebesar 0,77, rata-rata *recall* sebesar 0,74, dan rata-rata F1-score sebesar 0,76 menunjukkan bahwa secara keseluruhan,

performa model pada kedua kelas cukup baik. Namun demikian, temuan tersebut juga menunjukkan bahwa model masih dapat ditingkatkan untuk mengidentifikasi kelas minoritas. Model ini berarti bahwa prediksi yang dihasilkan harus lebih seimbang untuk setiap kelas.

```
=====
HASIL EVALUASI MODEL
=====
Accuracy : 0.9628
Precision: 0.5714
Recall   : 0.5
F1 Score : 0.5333
ROC AUC  : 0.9451

CLASSIFICATION REPORT
      precision    recall  f1-score   support

0         0.98      0.98      0.98         180
1         0.57      0.50      0.53           8

 accuracy          0.96         188
 macro avg         0.77         188
 weighted avg      0.96         188
```

Gambar 2. Klasifikasi XGBoots

Performa model XGBoost yang diperoleh menunjukkan bahwa algoritma ini mampu membedakan sampel *oil spill* dan *non-oil spill* dengan baik berdasarkan fitur hasil ekstraksi citra SAR. Kemampuan tersebut didukung oleh mekanisme *gradient boosting* yang memungkinkan model membangun pohon keputusan secara bertahap untuk memperbaiki kesalahan prediksi sebelumnya. Dalam penelitian ini, nilai *accuracy* yang tinggi menunjukkan bahwa sebagian besar sampel berhasil diklasifikasikan dengan benar. Namun demikian, pada kasus deteksi tumpahan minyak, nilai *recall* memiliki peran yang lebih penting karena menunjukkan kemampuan model dalam menemukan kejadian *oil spill* yang sebenarnya terdapat pada data. Nilai *recall* yang tinggi mengindikasikan bahwa risiko kegagalan deteksi dapat diminimalkan sehingga model lebih sesuai digunakan untuk mendukung proses pemantauan lingkungan laut.

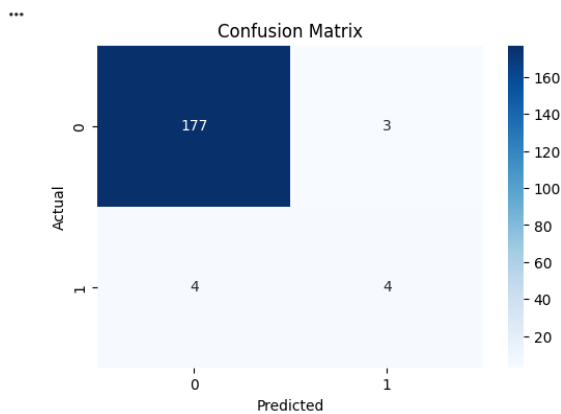
Performa model XGBoost menunjukkan kemampuan untuk membedakan sampel *oil*

spill dan sampel *non-oil spill* berdasarkan karakteristik hasil ekstraksi citra SAR. Kemampuan ini didukung oleh mekanisme peningkatan gradient, yang memungkinkan model secara bertahap membangun decision tree untuk memperbaiki kesalahan prediksi sebelumnya. Nilai akurasi tinggi penelitian ini menunjukkan bahwa sebagian besar sampel diklasifikasikan dengan benar. Namun, dalam kasus deteksi tumpahan minyak, nilai *recall* lebih penting karena menunjukkan kemampuan model untuk menemukan kejadian tumpahan *oil spill* yang sebenarnya terdapat pada data. Nilai *recall* yang tinggi menunjukkan bahwa risiko kegagalan deteksi dapat diminimalkan, yang berarti model lebih sesuai digunakan untuk mendukung proses pemantauan lingkungan laut.

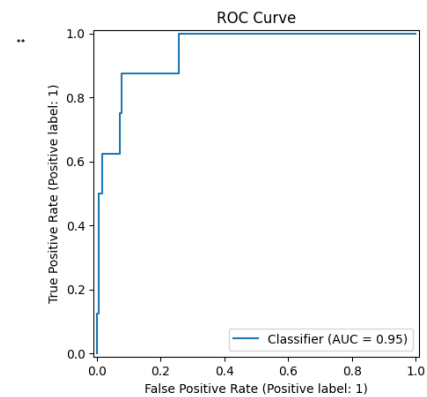
Penggunaan data *Synthetic Aperture Radar* (SAR) sebagai sumber informasi utama tidak mempengaruhi keberhasilan model dalam penelitian ini. Topouzelis (2008) menyatakan bahwa citra SAR adalah salah satu teknologi penginderaan jauh yang paling efektif untuk mendeteksi tumpahan oil spill. Hal ini dapat dilakukan baik pada siang maupun malam hari dan tidak terpengaruh oleh tutupan awan atau kondisi cuaca. Lapisan *oil spill* di permukaan laut dapat meredam gelombang kapiler, yang menghasilkan area gelap atau pembentukan gelap pada gambar SAR. Area gelap ini memungkinkan proses ekstraksi fitur dilakukan untuk membedakan area lepas *oil spill* dari area lepas *non-oil spill*. Akibatnya, data SAR yang digunakan dalam penelitian ini memiliki dasar ilmiah yang kuat dan telah banyak digunakan dalam penelitian sebelumnya yang mendeteksi tumpahan minyak.

Analisis *confusion matrix* menunjukkan bahwa sebagian besar sampel berhasil diklasifikasikan dengan benar oleh model. Dalam konteks operasional, jumlah *False Negative* (Gambar 4) menjadi indikator yang sangat penting karena menunjukkan banyaknya kejadian *oil spill* yang gagal terdeteksi. Berdasarkan hasil confusion matrix, model berhasil mengklasifikasikan 177 data kelas 0 dengan benar (Gambar 3) dan hanya melakukan kesalahan pada 3 data yang diprediksi sebagai kelas 1. Sementara itu, dari 8 data yang termasuk kelas 1, model mampu mengidentifikasi 4 data dengan benar, tetapi masih terdapat 4 data yang salah diklasifikasikan sebagai kelas 0. Hasil ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengenali kelas 0 yang jumlahnya lebih banyak, namun performanya dalam mendeteksi kelas 1 masih perlu ditingkatkan. Meskipun demikian, model tetap mampu menemukan sebagian data kelas 1 yang ada sehingga dapat digunakan untuk proses klasifikasi, dengan catatan bahwa masih terdapat beberapa data kelas 1 yang berpotensi tidak terdeteksi oleh model. Hal ini sejalan dengan nilai *accuracy* yang tinggi, tetapi nilai *recall* dan F1-score pada kelas 1 yang masih berada pada tingkat sedang. Semakin kecil nilai *False Negative*,

semakin baik kemampuan model dalam mengidentifikasi potensi tumpahan minyak. Menurut Topouzelis (2008), salah satu tantangan terbesar dalam deteksi oil spill adalah membedakan tumpahan minyak dengan fenomena laut lain yang memiliki tampilan serupa (*look-alikes*) seperti daerah berangin lemah, lapisan biologis, atau fenomena oseanografi lainnya. Oleh karena itu, kemampuan model dalam meminimalkan kesalahan klasifikasi menjadi aspek penting dalam pengembangan sistem deteksi otomatis.

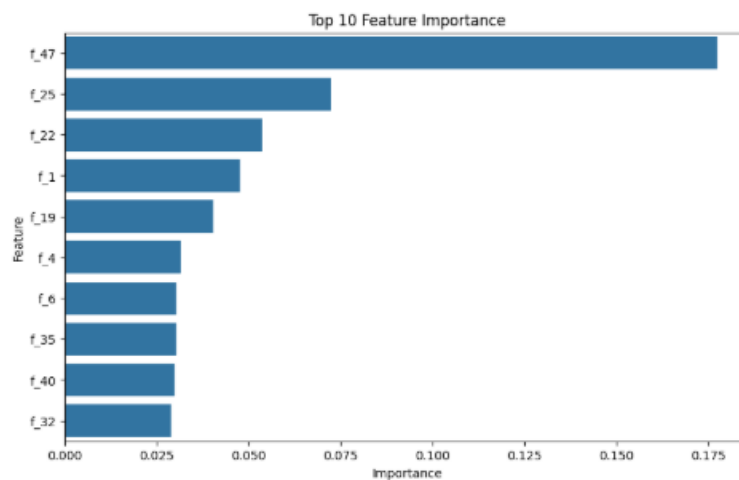


Gambar 3. Analisis *Confusion Matrix*



Gambar 4. Kurva ROC

Selain mengklasifikasikan *oil spill* dan *non-oil spill*, model juga menghasilkan nilai kemungkinan yang menunjukkan tingkat keyakinan model terhadap hasil prediksi. Nilai kemungkinan tersebut kemudian dikategorikan ke dalam beberapa kategori risiko: rendah, sedang, tinggi, dan sangat tinggi. Hasil klasifikasi risiko menunjukkan bahwa setiap *sample* dapat diberikan tingkat risiko yang sesuai dengan kemungkinan yang dihasilkan model. Seseorang dapat menggunakan data ini sebagai sinyal awal untuk memutuskan mana sampel yang memiliki risiko lebih tinggi yang harus diprioritaskan untuk dipantau.



Gambar 5. *Feature Performance*

Model yang dikembangkan tidak hanya menghasilkan klasifikasi tumpahan *oil spill* dan *non-oil spill*, tetapi juga menghasilkan nilai probabilitas yang digunakan untuk membentuk kategori risiko. Metode ini tidak hanya memungkinkan sistem untuk membuat keputusan biner, tetapi juga memungkinkan sistem untuk menunjukkan tingkat keyakinan model terhadap hasil prediksi yang dihasilkan. Sampel dengan probabilitas rendah dianggap risiko rendah. Sampel dengan probabilitas tinggi dianggap risiko tinggi atau sangat tinggi. Menurut Al-Ruzouq *et al.* (2020), *machine learning* dan penginderaan jauh telah meningkatkan proses deteksi serta sistem pemantauan, yang memungkinkan pengambilan keputusan yang lebih cepat dan efektif. Oleh karena itu, model yang dikembangkan dalam penelitian ini dapat digunakan untuk mendukung sistem peringatan dini berbasis data SAR.

```
*** HASIL DETEKSI DINI
      Probability_OilSpill Risk_Level
0      0.000730      Low
1      0.000744      Low
2      0.001209      Low
3      0.007734      Low
4      0.000525      Low
5      0.000559      Low
6      0.013909      Low
7      0.000698      Low
8      0.001441      Low
9      0.000713      Low
10     0.000258      Low
11     0.052272      Low
12     0.001588      Low
13     0.001336      Low
14     0.024594      Low
15     0.002934      Low
16     0.049711      Low
17     0.000440      Low
18     0.000364      Low
19     0.001213      Low
```

Gambar 6. Hasil Deteksi Dini

KESIMPULAN

Penelitian ini berhasil mengembangkan model deteksi dini potensi tumpahan minyak menggunakan algoritma XGBoost berdasarkan fitur hasil ekstraksi citra *Synthetic Aperture Radar* (SAR). Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan sampel *oil spill* dan *non-oil spill* dengan performa yang baik, yang ditunjukkan melalui nilai evaluasi seperti *accuracy*, *precision*, *recall*, *F1-score*, dan *ROC-AUC*. Penerapan metode SMOTE juga terbukti membantu mengatasi ketidakseimbangan data sehingga meningkatkan kemampuan model dalam mengenali kelas *oil spill*. Selain menghasilkan klasifikasi, model mampu memberikan nilai probabilitas yang kemudian dikategorikan ke dalam beberapa tingkat risiko, sehingga berpotensi digunakan sebagai komponen pendukung sistem deteksi dini tumpahan

minyak berbasis data SAR. Meskipun demikian, keterbatasan dataset yang tidak memiliki informasi spasial dan temporal menyebabkan model belum dapat digunakan untuk menentukan lokasi maupun sebaran tumpahan minyak secara langsung. Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan data SAR yang memiliki informasi geografis dan faktor lingkungan lainnya guna meningkatkan kemampuan sistem dalam mendukung pemantauan dan pengelolaan pencemaran laut secara lebih komprehensif.

DAFTAR PUSTAKA

- Al-Ruzouq, R., Gibril, M. B. A., Shanableh, A., Kais, A., Hamed, O., Al-Ali, E., Khalil, M. A., Idris, A., & Al-Ruzouq, M. (2020). Oil spill detection using machine learning and deep learning: A review. *Remote Sensing*, *12*(20), 3338. <https://doi.org/10.3390/rs12203338>
- Bianchi, F. M., Espeseth, M. M., & Borch, N. (2020). Large-scale detection and categorization of oil spills from SAR images with deep learning. *Remote Sensing*, *12*(14), 2260. <https://doi.org/10.3390/rs12142260>
- Conceição, M. R. A., Braga, A. P., & Silva, C. R. (2021). Oil spill detection in Sentinel-1 SAR images using machine learning techniques. *Remote Sensing*, *13*(11), 2044. <https://doi.org/10.3390/rs13112044>
- Topouzelis, K. (2008). Oil spill detection by SAR images: Dark formation detection, feature extraction and classification algorithms. *Sensors*, *8*(10), 6642–6659. <https://doi.org/10.3390/s8106642>
- Zeng, K., & Wang, Y. (2020). Deep learning-based oil spill detection from synthetic aperture radar images: A review and future perspectives. *Remote Sensing*, *12*(20), 3338. <https://doi.org/10.3390/rs12061015>
- Marsisno, W. (2025). Oil spill detection and verification in northern Bintan. *ZERO: Jurnal Sains, Matematika dan Terapan*, *9*(2), 555–564. <https://doi.org/10.30829/zero.v9i2.25942>
- Prastyani, R., & Basith, A. (2019). DETEKSI TUMPAHAN MINYAK DI SELAT MAKASSAR DENGAN PENGINDERAAN JAUH SENSOR AKTIF DAN PASIF. *Jurnal Geodesi dan Geomatika*, *2*(1), 89-90. <https://ejournal2.undip.ac.id/index.php/elipsoida/article/view/4864/2899>
- Suniada, K. I. (2021). PEMANFAATAN DATA CITRA SATELIT SENTINEL-1 UNTUK PEMANTAUAN SEBARAN TUMPAHAN MINYAK DI WPP 713. *JURNAL KELAUTAN NASIONAL*, *16*(1), 16-17. <https://ejournal-balitbang.kkp.go.id/index.php/jkn/article/view/9293/pdf>
- Supriyanto, B. F., Hanifa, S., Nurmawati, I., & Ramdani, F. (2025). Pengaruh Tumpahan Minyak di Lautan Terhadap Penyakit Di Masyarakat Pesisir Pantai. *Journal of Artificial Intelligence and Digital Business (RIGGS)*, *4*(2), 2270-2278. <https://doi.org/10.31004/riggs.v4i2.928>
- Sutoyo, E., & Fadlurranman, M. A. (2020). Penerapan SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Television Advertisement Performance Rating Menggunakan Artificial Neural Network. *Jurnal Edukasi dan Penelitian Informatika*, *6*(3), 379-380. <https://doi.org/10.26418/jp.v6i3.42896>