

## IDENTIFIKASI BANJIR ROB MENGGUNAKAN METODE KLASIFIKASI DENGAN MODEL RANDOM FOREST DAN DECISION TREE DI PELABUHAN SURABAYA TAHUN 2021-2023

<sup>1)</sup>Kiki Syalasyatun Masfufah, <sup>2)</sup> Kartika Dwi Indra Setyaningrum, <sup>3)</sup> Endah Rahmawati, <sup>4)</sup>Ady Hermanto

<sup>1)</sup> Program Studi Fisika, FMIPA, Universitas Negeri Surabaya, email: [kiki.22041@mhs.unesa.ac.id](mailto:kiki.22041@mhs.unesa.ac.id)

<sup>2)</sup> Program Studi Fisika, FMIPA, Universitas Negeri Surabaya, email: [kartika.22044@mhs.unesa.ac.id](mailto:kartika.22044@mhs.unesa.ac.id)

<sup>3)</sup> Program Studi Fisika, FMIPA, Universitas Negeri Surabaya, email: [endahrahmawati@unesa.ac.id](mailto:endahrahmawati@unesa.ac.id)

<sup>4)</sup> BMKG, Tanjung Perak Surabaya, email: [ady.hermanto@bmkg.go.id](mailto:ady.hermanto@bmkg.go.id)

### Abstrak

Banjir rob merupakan salah satu jenis bencana hidrometeorologi yang kerap terjadi di wilayah pesisir, terutama akibat kombinasi antara pasang laut tinggi, fase bulan tertentu, dan curah hujan ekstrem. Penelitian ini bertujuan untuk mengidentifikasi potensi banjir rob di kawasan Pelabuhan Surabaya dengan menggunakan algoritma klasifikasi *Random Forest* dan *Decision Tree*. Data yang digunakan meliputi parameter hidrometeorologi seperti fase bulan, pasang surut air laut, dan curah hujan, yang diperoleh dari Stasiun Meteorologi Maritim Tanjung Perak Surabaya. Sebelum dimodelkan, data melalui tahapan *pre-processing* berupa normalisasi dan pembobotan untuk menyeragamkan skala antar variabel serta menilai tingkat kontribusi masing-masing parameter. Model dikembangkan dengan pembagian data 80% pelatihan dan 20% pengujian. Hasil evaluasi *Random Forest* menunjukkan akurasi sebesar 99,96% dan *Decision Tree* sebesar 99,94% dengan tingkat kesalahan yang sangat rendah. Analisis *feature importance* menunjukkan bahwa fase bulan dan pasang surut merupakan faktor dominan dalam prediksi banjir rob, sedangkan curah hujan memiliki pengaruh minimal. Temuan ini membuktikan bahwa *Random Forest* merupakan metode yang efektif dan andal untuk klasifikasi banjir rob serta memiliki potensi untuk diimplementasikan dalam sistem peringatan dini. Penelitian ini juga merekomendasikan integrasi data geografis, seperti informasi kerentanan tanah, morfologi wilayah, dan elevasi permukaan untuk meningkatkan akurasi dan generalisasi model di masa mendatang.

**Kata Kunci:** banjir rob, *Random Forest*, *Decision Tree*, klasifikasi, hidrometeorologi

### Abstract

Tidal flooding is a type of hydrometeorological disaster that often occurs in coastal areas, mainly due to a combination of high tides, certain lunar phases, and extreme rainfall. This study aims to identify the potential for tidal flooding in the Surabaya Port area using *Random Forest* and *Decision Tree* classification algorithms. The data used include hydrometeorological parameters such as lunar phases, tides, and rainfall, obtained from the Tanjung Perak Maritime Meteorology Station in Surabaya. Before being modeled, the data went through *pre-processing* stages such as normalization and weighting to standardize the scale between variables and assess the level of contribution of each parameter. The model was developed by dividing the data into 80% training and 20% testing. The evaluation results of *Random Forest* showed an accuracy of 99.96% and *Decision Tree* at 99.94% with a very low error rate. Feature importance analysis showed that lunar phases and tides are the dominant factors in predicting tidal flooding, while rainfall has a minimal influence. These findings prove that *Random Forest* is an effective and reliable method for tidal flood classification and has the potential to be implemented in early warning systems. This study also recommends the integration of geographic data, such as soil vulnerability information, regional morphology, and surface elevation to improve the accuracy and generalization of future models.

**Keywords:** Tidal flooding, *Random Forest*, *Decision Tree*, Classification, Hydrometeorology

## I. PENDAHULUAN

Banjir rob merupakan salah satu bencana hidrometeorologi yang terjadi akibat meluapnya air laut ke daratan, terutama di wilayah pesisir dengan elevasi rendah. Fenomena ini dipengaruhi oleh interaksi berbagai faktor fisik, seperti pasang surut laut, fase bulan, tekanan atmosfer, serta curah hujan (Vousdoukas *et al.*, 2018; Nicholls *et al.*, 2018). Kombinasi pasang laut tinggi dengan kondisi hidrometeorologi tertentu dapat meningkatkan tinggi muka air laut dan memperbesar peluang terjadinya genangan di wilayah pesisir (Wahl *et al.*, 2015; Bevacqua *et al.* 2019).

Pelabuhan Tanjung Perak Surabaya merupakan salah satu pelabuhan terbesar dan tersibuk di Jawa bagian timur yang memiliki tingkat kerentanan tinggi terhadap banjir rob. Kejadian banjir rob di kawasan ini berpotensi menimbulkan kerusakan infrastruktur, gangguan distribusi logistik, peningkatan biaya operasional, serta ancaman keselamatan bagi pekerja dan masyarakat sekitar. Oleh karena itu, diperlukan sistem identifikasi potensi banjir rob yang akurat sebagai dasar mitigasi dan perencanaan operasional pelabuhan.

Secara fisik, pasang surut laut merupakan fenomena periodik naik turunnya permukaan laut yang dipengaruhi oleh gaya gravitasi Bulan dan Matahari. Variasi tinggi muka air laut dapat dimodelkan secara matematis menggunakan pendekatan harmonik sederhana sebagai berikut (Pugh, 1987):

$$\eta_{\text{tide}}(t) = A \sin(\omega t + \phi) \quad (1)$$

dengan  $\eta_{\text{tide}}(t)$  menyatakan tinggi muka air laut pada waktu  $t$ ,  $A$  adalah amplitudo pasang,  $\omega$  adalah frekuensi sudut pasang surut, dan  $\phi$  merupakan fase awal. Persamaan ini menunjukkan bahwa tinggi muka air laut berfluktuasi secara periodik, dengan nilai maksimum yang umumnya terjadi pada fase bulan baru dan purnama (spring tide). Variasi amplitudo dan fase pasang surut terbukti berperan penting dalam menentukan ketinggian muka air laut serta durasi genangan banjir pesisir (Haigh *et al.*, 2025; Phung *et al.*, 2021). Selain pasang surut, curah hujan berperan sebagai faktor pendukung yang dapat memperparah genangan banjir rob. Secara sederhana, kontribusi curah hujan terhadap peningkatan tinggi muka air di daratan dapat dinyatakan melalui hubungan limpasan berikut (Ward & Robinson, 2000):

$$\Delta\eta_{\text{hujan}} = \beta R \quad (2)$$

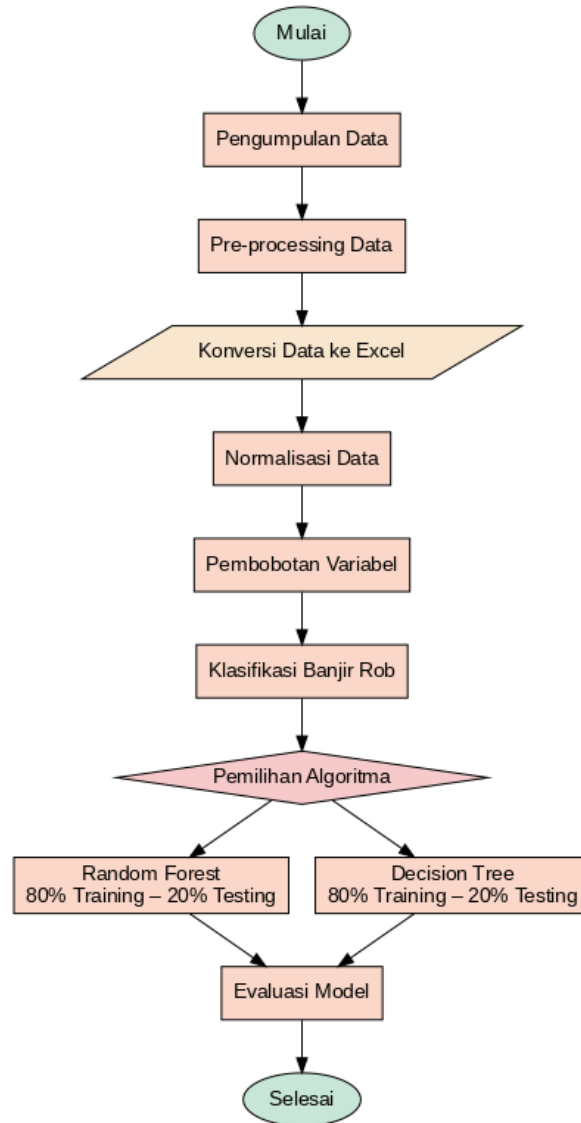
dengan  $\beta$  sebagai koefisien limpasan dan  $R$  adalah intensitas curah hujan. Pada kawasan perkotaan dan pelabuhan dengan dominasi permukaan kedap air, nilai  $\beta$  relatif besar, sehingga sebagian besar hujan langsung menjadi limpasan permukaan. Dalam konteks banjir rob, curah hujan tidak berperan sebagai penyebab utama, tetapi bertindak sebagai faktor penguat yang meningkatkan intensitas dan durasi genangan ketika terjadi bersamaan dengan pasang laut maksimum (Setiawan *et al.*, 2018).

Meskipun peran pasang surut dan curah hujan terhadap banjir rob telah banyak dibahas, pemanfaatan metode machine learning untuk mengklasifikasikan potensi banjir rob berbasis data hidrometeorologi historis di kawasan pelabuhan masih relatif terbatas. Padahal, algoritma Random Forest dan Decision Tree telah terbukti efektif dalam memodelkan fenomena hidrometeorologi yang bersifat nonlinier dan kompleks (Breiman, 2001; Quinlan, 1986; Mobley *et al.*, 2021). Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi potensi banjir rob di Pelabuhan Tanjung Perak Surabaya menggunakan algoritma Random Forest dan Decision Tree dengan parameter fase bulan, pasang surut, dan curah hujan. Penelitian ini diharapkan dapat mendukung pengembangan sistem identifikasi dan peringatan dini banjir rob yang lebih presisi di wilayah pelabuhan.

## II. METODE

## A. Rancangan Penelitian

Penelitian ini dirancang untuk mengidentifikasi potensi banjir rob di Pelabuhan Tanjung Perak Surabaya dengan memanfaatkan data hidrometeorologi dari BMKG serta metode klasifikasi *Random Forest* dan *Decision Tree*. Tahapan penelitian dimulai dari pengumpulan data, pengolahan awal (pre-processing), penentuan kategori potensi banjir rob, hingga penerapan model machine learning. Pre-processing dilakukan untuk memastikan data dalam kondisi bersih, seragam, dan memiliki skala yang sama, sehingga proses klasifikasi dapat berjalan optimal. Alur lengkap tahapan penelitian ini digambarkan pada diagram berikut.



Gambar 1. Flowchart Sistem Penelitian

Berdasarkan diagram alur tersebut, penelitian diawali dengan pengumpulan data hidrometeorologi yang meliputi **pasang surut air laut**, **fase bulan (fasemoon)**, dan **curah hujan**, yang merupakan parameter utama dalam kejadian banjir rob. Tahap pengolahan awal data dilakukan untuk memastikan kualitas data sebelum digunakan dalam proses klasifikasi. Tahap ini mencakup upaya pembersihan data, penyamaan skala antarparameter, serta penguatan kontribusi variabel dominan agar proses pembelajaran model dapat berjalan optimal dan menghasilkan klasifikasi yang stabil (Han, Kamber, & Pei, 2012). Klasifikasi data berdasarkan nilai gabungan hasil normalisasi dan pembobotan menjadi tiga kategori potensi banjir rob, yaitu 0-0,4 dikategorikan sebagai tidak terjadi banjir rob, >0,4-≤0,6 sebagai banjir rob ringan, dan >0,6-1 sebagai banjir rob tinggi.

Hasil klasifikasi ini kemudian digunakan sebagai input untuk membangun model klasifikasi berbasis *Random Forest* dan *Decision Tree*. Pembagian ini bertujuan untuk melatih model dalam mengenali pola hubungan antarparameter hidrometeorologi terhadap kejadian banjir rob, serta menguji kemampuannya dalam mengidentifikasi kategori banjir rob pada data baru yang belum pernah digunakan sebelumnya. Dengan pendekatan ini, model diharapkan mampu memberikan hasil identifikasi yang akurat, konsisten, dan dapat diandalkan sebagai pendukung sistem peringatan dini dan mitigasi risiko banjir rob di wilayah pesisir.

**Tabel 1.** Data Penelitian

Data set	Rentang data	Sumber data
Fase bulan	0 hari – 29,5 hari	Timeanddate moon phase <a href="https://timeanddate.com">timeanddate.com</a>
Pasut	-120 cm – 160 cm	DISHIDROS <a href="https://www.pushidrosal.id/">https://www.pushidrosal.id/</a>
Curah hujan	0 mm – 102,4 mm	BMKG Stasiun Perak 1 Surabaya

Dalam penelitian ini, data berjumlah 26.275 lalu data yang digunakan dibagi menjadi dua bagian, yaitu 21.020 data training sebesar 80% dan 5.255 data uji sebesar 20%. Pembagian ini dilakukan secara acak (random) untuk memastikan bahwa model tidak hanya mengenali pola dari data tertentu saja, melainkan dapat melakukan generalisasi terhadap data baru yang belum pernah dilihat sebelumnya. Pada tahap training, model *Random Forest* akan mempelajari pola hubungan antara variabel input seperti fasemoon, pasang surut, dan curah hujan dengan tingkat potensi banjir rob. Selanjutnya, pada tahap pengujian, model akan diuji menggunakan data uji untuk mengevaluasi performanya dalam mengklasifikasikan data yang tidak dikenal. Dengan pembagian data yang proporsional dan acak, model diharapkan mampu memberikan hasil klasifikasi yang akurat dan tidak bias, serta menjadi dasar yang kuat dalam sistem identifikasi banjir rob secara otomatis.

## **B. Teknik Pengumpulan Data**

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), khususnya Stasiun Meteorologi Maritim Perak Surabaya. Parameter yang dikumpulkan meliputi pasang surut air laut, fase bulan, dan curah hujan, yang seluruhnya bersumber dari BMKG. Penggunaan satu sumber data yang konsisten dilakukan untuk menghindari perbedaan datum, resolusi waktu, dan standar pengukuran yang berpotensi menimbulkan bias dalam analisis hidrometeorologi.

Data pasang surut dipilih dalam bentuk data numerik per jam yang direferensikan terhadap Mean Sea Level (MSL), karena resolusi temporal ini mampu merepresentasikan fluktuasi pasang surut secara detail dan menangkap kondisi pasang maksimum yang relevan terhadap kejadian banjir rob. Periode data ditentukan dengan menyesuaikan ketersediaan data curah hujan dan laporan kejadian banjir rob dari BMKG, sehingga seluruh parameter dianalisis pada rentang waktu yang sama untuk menjaga keterkaitan temporal antarvariabel. Koreksi data pasut dilakukan secara sederhana melalui penyelarasan waktu (time alignment) dan pengecekan konsistensi nilai antara data numerik dan grafik pasang surut, serta penyaringan nilai ekstrem yang tidak merepresentasikan kondisi pasang surut astronomis, guna meminimalkan kesalahan digitasi dan anomali data.

## **C. Teknik Pengolahan Data**

Teknik pengolahan data dalam penelitian ini meliputi proses pre-processing yang terdiri atas normalisasi, pembobotan variabel, dan klasifikasi awal. Normalisasi dilakukan dengan mengubah seluruh variabel ke dalam skala 0–1 untuk menghilangkan perbedaan skala antarparameter, di mana nilai 1 menunjukkan kondisi dengan potensi banjir rob tertinggi. Selanjutnya, dilakukan pembobotan variabel untuk mencerminkan tingkat pengaruh relatif masing-masing parameter terhadap potensi banjir rob, sebagai pendekatan umum dalam pemodelan prediktif guna meningkatkan akurasi hasil klasifikasi (Guyon & Elisseeff, 2003). Dalam penelitian ini, bobot fase bulan ditetapkan sebesar 0,55, pasang surut 0,35, dan curah



hujan 0,10, berdasarkan kajian literatur yang menunjukkan dominasi pasut dan fase bulan dalam kejadian banjir rob.

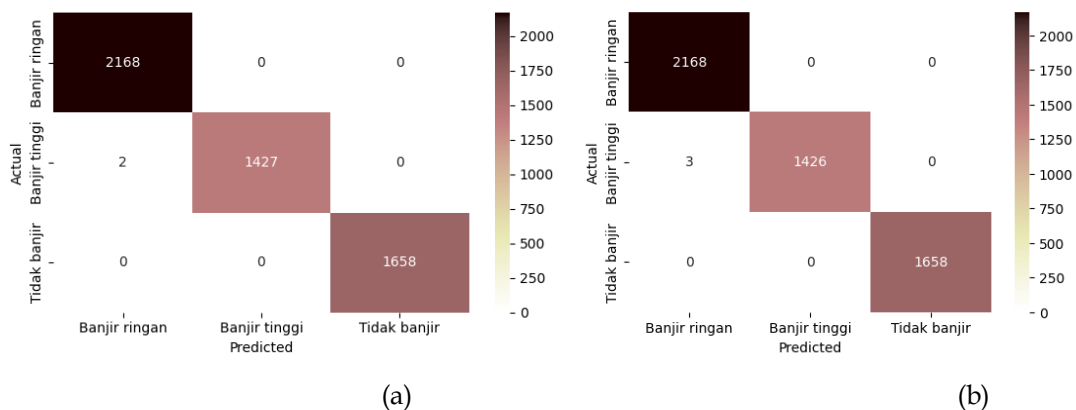
Untuk mendukung analisis spasial genangan banjir rob, digunakan Digital Elevation Model Nasional (DEMNAS) yang dikembangkan oleh Badan Informasi Geospasial (BIG). DEMNAS dibangun menggunakan metode Triangulated Irregular Network (TIN) dengan resolusi spasial sekitar 8 m dan memiliki ketelitian vertikal dengan nilai RMSE berkisar 0,8–1,5 m, bergantung pada kondisi topografi wilayah. Ketidakpastian elevasi ini menjadi salah satu keterbatasan penelitian dan berpotensi memengaruhi estimasi luas genangan, khususnya di wilayah pesisir dengan topografi sangat landai (Biggs *et al.*, 2013).

### III. HASIL DAN PEMBAHASAN

#### A. Hasil

Berdasarkan hasil pengujian terhadap beberapa konfigurasi, diperoleh bahwa model *Random Forest* dengan parameter *max\_depth* = None, *min\_samples\_split* = 2, dan *n\_estimators* = 100 menghasilkan akurasi tertinggi sebesar 0,9996 (99,96%). Nilai ini merupakan yang terbaik dibandingkan parameter lainnya. Selanjutnya pada model *Decision Tree* akurasi terbaik diperoleh ketika menggunakan parameter *criterion* = *entropy*, *max\_depth* = 100, *min\_samples\_split* = 4, dan *min\_samples\_leaf* = 1. Penggunaan *criterion* = *entropy* berarti model membagi data berdasarkan informasi yang memberikan *information gain* tertinggi, sehingga pemisahan antar kelas menjadi lebih optimal.

Pada penelitian ini dikembangkan model klasifikasi potensi banjir rob dengan menggunakan algoritma *Random Forest* dan *Decision Tree*. Hasil pengujian menunjukkan bahwa algoritma *Random Forest* memberikan tingkat akurasi sebesar 99,96%, sedangkan *Decision Tree* mencapai 99,94%. Evaluasi kinerja model dilakukan dengan menggunakan confusion matrix pada tiga kelas utama, yaitu banjir ringan, banjir tinggi, dan tidak banjir.



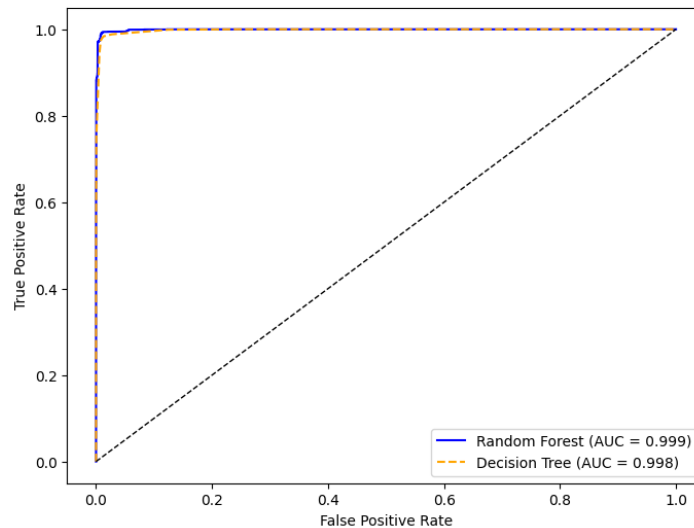
Gambar 2. (a) Confusion Matrix *Random Forest* , (b) Confusion Matrix *Decision Tree*

Berdasarkan Confusion Matrix diatas dihasilkan bahwa model *Random Forest* seluruh data pada kelas banjir ringan (2.168 data) dan tidak banjir (1.658 data) berhasil diklasifikasikan dengan benar tanpa adanya kesalahan. Pada kelas banjir tinggi, dari total 1.429 data terdapat 2 data yang salah klasifikasi. Hal ini menunjukkan bahwa *Random Forest* mampu menghasilkan prediksi yang lebih stabil dan akurat karena menggunakan gabungan banyak pohon keputusan dalam proses klasifikasi.

Sementara itu, pada model *Decision Tree* hasil klasifikasi juga menunjukkan tingkat akurasi yang sangat tinggi. Sama halnya dengan *Random Forest*, seluruh data pada kelas *banjir ringan* dan *tidak banjir* berhasil diprediksi dengan benar. Namun, pada kelas *banjir tinggi* terdapat 3 data yang salah klasifikasi dari total 1.429 data. Hal ini menunjukkan bahwa meskipun *Decision Tree* efektif dalam mempelajari pola data, model ini lebih rentan terhadap kesalahan pada data yang berada di batas antar kelas dibandingkan dengan *Random Forest*.

Secara keseluruhan, dapat disimpulkan bahwa *Random Forest* memiliki performa yang lebih baik dibandingkan *Decision Tree*, dengan jumlah kesalahan prediksi lebih sedikit, khususnya pada kategori banjir

tinggi. Dengan demikian, model ensemble seperti *Random Forest* lebih unggul dalam mengklasifikasikan potensi banjir rob karena kemampuannya mengurangi overfitting dan meningkatkan stabilitas prediksi.



**Gambar 4.** ROC Curve *Random Forest* vs *Decision Tree*

Gambar ROC Curve menunjukkan bahwa model *Random Forest* dan *Decision Tree* sama-sama memiliki kinerja yang sangat baik dalam membedakan antara kelas banjir rob dan tidak banjir rob. Nilai AUC *Random Forest* sebesar 0,999 dan AUC *Decision Tree* sebesar 0,998, keduanya mendekati nilai maksimum (1,0). Hal ini mengindikasikan bahwa kedua model mampu mengklasifikasikan data dengan tingkat akurasi yang sangat tinggi, meskipun *Random Forest* sedikit lebih unggul dibandingkan *Decision Tree*.

Hasil analisis *feature importance* menunjukkan bahwa variabel fase bulan memiliki kontribusi terbesar sebesar 0,75, diikuti oleh pasang surut sebesar 0,22, sedangkan curah hujan hanya mendekati nol. Dominasi fase bulan dapat dijelaskan karena pada fase bulan baru dan purnama terjadi pasang maksimum akibat gaya tarik gravitasi bulan dan matahari yang bekerja bersamaan, sehingga meningkatkan ketinggian muka laut. Pasang surut menjadi faktor kedua yang signifikan karena secara langsung menentukan terjadinya limpasan air laut ke daratan ketika melebihi ambang topografi pesisir. Sebaliknya, curah hujan tidak berpengaruh signifikan terhadap banjir rob karena perannya lebih dominan pada banjir genangan atau luapan sungai.

## B. Pembahasan

Hasil pengujian menunjukkan bahwa model *Random Forest* memiliki performa yang lebih baik dibandingkan *Decision Tree* dalam mengklasifikasikan potensi banjir rob. Keunggulan ini dapat dijelaskan oleh karakteristik *Random Forest* sebagai algoritma *ensemble* yang menggabungkan banyak pohon keputusan, sehingga mampu mereduksi kesalahan prediksi yang bersifat acak dan meningkatkan stabilitas model. Mekanisme *majority voting* memungkinkan *Random Forest* menangani data dengan karakteristik ambiguitas antar kelas secara lebih efektif dibandingkan *Decision Tree* yang hanya mengandalkan satu struktur pohon keputusan.

Kesalahan klasifikasi yang muncul pada kedua model terutama terjadi pada batas antara kelas banjir rob tinggi dan banjir rob ringan. Kondisi ini menunjukkan bahwa kedua kelas tersebut memiliki karakteristik hidrometeorologi yang saling berdekatan, misalnya nilai pasang surut yang relatif tinggi tanpa disertai curah hujan ekstrem, atau fase bulan yang mendukung pasang maksimum dengan amplitudo yang tidak jauh berbeda. Pada kondisi seperti ini, pemisahan kelas menjadi kurang tegas secara numerik, sehingga meningkatkan potensi salah klasifikasi, khususnya pada model *Decision Tree* yang memiliki aturan pemisahan lebih kaku.

Dominasi variabel fase bulan dalam analisis *feature importance* menunjukkan bahwa siklus pasang maksimum (*spring tide*) merupakan faktor utama dalam menentukan potensi terjadinya banjir rob di wilayah pesisir Surabaya. Secara fisik, kondisi ini terjadi karena superposisi gaya gravitasi Bulan dan Matahari pada

fase bulan baru dan purnama yang meningkatkan tinggi muka air laut. Pasang surut menjadi faktor pendukung yang memperkuat peluang limpasan air laut ke daratan ketika ketinggian muka laut melampaui elevasi topografi pesisir. Sebaliknya, curah hujan memiliki pengaruh yang lebih kecil terhadap banjir rob karena perannya lebih dominan pada banjir genangan atau luapan sungai. Temuan ini sejalan dengan studi (Marfai & King, 2021) yang memperkirakan perluasan genangan banjir rob di Jawa akibat kenaikan muka laut, sehingga memperkuat pentingnya penggunaan model prediktif untuk memetakan area terdampak. Hal tersebut juga konsisten dengan (Shen *et al.*, 2022) yang menekankan kompleksitas interaksi non-linear antara pasang surut, storm surge, dan aliran sungai, yang lebih tepat dimodelkan dengan algoritma ensemble seperti *Random Forest* dibandingkan *Decision Tree* tunggal.

Sebagai pembandingan terhadap kondisi aktual, hasil klasifikasi model dibandingkan dengan kejadian banjir rob yang dilaporkan oleh BMKG pada 11 Agustus 2022 di wilayah pesisir Surabaya. Pada tanggal tersebut, BMKG melaporkan terjadinya banjir rob yang menggenangi kawasan pelabuhan dan sekitarnya. Hasil prediksi model pada periode yang sama menunjukkan kategori banjir rob tinggi, baik pada model *Random Forest* maupun *Decision Tree*. Kesesuaian ini menunjukkan bahwa model mampu merepresentasikan kondisi hidrometeorologi ekstrem yang memicu kejadian banjir rob.

Secara keseluruhan, temuan penelitian ini konsisten dengan studi-studi sebelumnya yang menunjukkan bahwa algoritma ensemble seperti *Random Forest* lebih unggul dalam memodelkan hubungan non-linear dan menangani ambiguitas antar kelas dibandingkan *Decision Tree* tunggal (Cutler *et al.*, 2007). Keunggulan tersebut menjadikan *Random Forest* lebih sesuai untuk aplikasi prediksi operasional banjir rob yang menuntut akurasi dan stabilitas tinggi, sementara *Decision Tree* tetap memiliki peran penting sebagai model yang bersifat interpretatif untuk memahami struktur keputusan serta kontribusi masing-masing variabel hidrometeorologi.

Secara keseluruhan, temuan ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa algoritma *ensemble* seperti *Random Forest* lebih unggul dalam menangani hubungan non-linear dan ambiguitas antar kelas dibandingkan *Decision Tree* tunggal. Dengan demikian, *Random Forest* dinilai lebih sesuai untuk aplikasi prediksi operasional banjir rob, sedangkan *Decision Tree* tetap bermanfaat sebagai model interpretatif untuk memahami struktur keputusan dan kontribusi variabel hidrometeorologi.

#### IV. PENUTUP

##### A. Simpulan

Berdasarkan hasil penelitian yang telah diperoleh, dapat disimpulkan bahwa Penentuan parameter hidrometeorologi dalam mengidentifikasi potensi terjadinya banjir rob memerlukan tahapan *pre-processing* data yang terstruktur dan sistematis. Proses ini mencakup normalisasi untuk menyeragamkan skala antar variabel, pembobotan guna menilai tingkat signifikansi relatif setiap parameter, serta klasifikasi output untuk membedakan antara kondisi yang berpotensi menimbulkan banjir rob dan yang tidak. Algoritma *Random Forest* dan *Decision Tree* digunakan dalam penelitian ini, di mana *Decision Tree* memberikan gambaran aturan klasifikasi secara sederhana dan interpretatif, sementara *Random Forest* terbukti lebih unggul dengan akurasi mencapai 99,96%. Keunggulan *Random Forest* juga terlihat pada analisis *feature importance*, di mana fase bulan (0,75) dan pasang surut (0,22) berperan dominan dalam memengaruhi potensi banjir rob, sedangkan curah hujan berkontribusi sangat kecil.

Penemuan ini menunjukkan bahwa *Random Forest* tidak hanya efektif sebagai metode klasifikasi, tetapi juga berpotensi untuk diterapkan ke dalam sistem peringatan dini banjir rob. Dengan penerapan yang tepat, model ini dapat menjadi dasar dalam pengambilan keputusan yang lebih cerdas dan responsif, khususnya dalam merancang strategi mitigasi risiko bencana di wilayah pesisir yang rentan terhadap dampak perubahan iklim dan kenaikan muka air laut.

##### B. Saran

Berdasarkan hasil penelitian yang telah diperoleh, disarankan agar studi selanjutnya mengintegrasikan data geografis, seperti informasi kerentanan tanah, morfologi wilayah, dan elevasi permukaan, guna

meningkatkan akurasi dalam mengidentifikasi area berisiko tinggi terhadap banjir rob. Selain itu, melakukan pengembangan model dengan data yang lebih luas dan variatif agar penelitian terkait banjir rob semakin akurat dan aplikatif. Dengan mengombinasikan pendekatan spasial dan algoritma klasifikasi yang lebih adaptif, diharapkan penelitian mendatang dapat menghasilkan sistem prediksi banjir rob yang lebih akurat, komprehensif, dan aplikatif dalam mendukung kebijakan mitigasi bencana serta sistem peringatan dini di wilayah pesisir.

### Acknowledgment

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Maritim Tanjung Perak Surabaya yang telah menyediakan data hidrometeorologi dan pasang surut yang menjadi dasar utama dalam penelitian ini. Tidak lupa penulis menyampaikan apresiasi kepada dosen pembimbing serta rekan-rekan yang senantiasa memberikan arahan, motivasi, dan masukan berharga sehingga penelitian ini dapat terselesaikan dengan baik.

### DAFTAR PUSTAKA

- Bevacqua, E., Vousdoukas, M. I., Shepherd, T. G., & Vrac, M. (2019). *Future changes in extreme storm surges along Europe*. *Geophysical Research Letters*, 46(3), 1344–1352. <https://doi.org/10.1029/2018GL081438>
- Biggs, J., Wright, T. J., Lu, Z., & Parsons, B. (2013). *Multi-interferogram method for measuring interseismic deformation: Denali Fault, Alaska*. *Geophysical Journal International*, 194(2), 753–760. <https://doi.org/10.1093/gji/ggt132>
- Breiman, L. (2001). *Random forests*. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). *Random forests for classification in ecology*. *Ecology*, 88(11), 2783–2792. <https://doi.org/10.1890/07-0539.1>
- Guyon, I., & Elisseeff, A. (2003). *An introduction to variable and feature selection*. *Journal of Machine Learning Research*, 3, 1157–1182. <https://www.jmlr.org/papers/v3/guyon03a.html>
- Haigh, I. D., Wahl, T., Rohling, E. J., Price, R. M., Pattiaratchi, C. B., Calafat, F. M., & Dangendorf, S. (2025). *Timescales for detecting future changes in sea level extremes*. *Nature Climate Change*, 15, 110–118. <https://doi.org/10.1038/s41558-024-02090-2>
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Marfai, M. A., & King, L. (2021). *Tidal inundation mapping under enhanced land subsidence in Semarang, Central Java Indonesia*. *Natural Hazards*, 105, 259–281. <https://doi.org/10.1007/s11069-020-04338-5>
- Mobley, R., Kiani, M., & Smith, J. (2021). *Machine learning approaches for flood susceptibility mapping*. *Journal of Hydrology*, 603, 127042. <https://doi.org/10.1016/j.jhydrol.2021.127042>
- Nicholls, R. J., Hanson, S., Lowe, J. A., Warrick, R. A., Lu, X., & Long, A. J. (2018). *Sea-level rise and its possible impacts given a “beyond 4°C world” in the twenty-first century*. *Philosophical Transactions of the Royal Society A*, 369(1934), 161–181. <https://doi.org/10.1098/rsta.2010.0291>
- Phung, D., Van, T., & Kim, S. (2021). *Assessment of tidal flooding using harmonic analysis*. *Ocean Engineering*, 235, 109400. <https://doi.org/10.1016/j.oceaneng.2021.109400>
- Pugh, D. T. (1987). *Tides, surges and mean sea-level*. John Wiley & Sons.
- Quinlan, J. R. (1986). *Induction of decision trees*. *Machine Learning*, 1(1), 81–106. <https://doi.org/10.1007/BF00116251>
- Setiawan, H., Nugroho, S. P., & Suryadi, F. X. (2018). *Analisis banjir rob di wilayah pesisir Indonesia*. *Jurnal Teknik Hidraulik*, 9(2), 85–94.
- Shen, Y., Zhang, X., & Wang, Y. (2022). *Nonlinear interactions between tides, storm surge, and river discharge*. *Journal of Hydrology*, 610, 127906. <https://doi.org/10.1016/j.jhydrol.2022.127906>
- Vousdoukas, M. I., Mentaschi, L., Feyen, L., & Voukouvalas, E. (2018). *Extreme sea levels on the rise along Europe’s coasts*. *Earth’s Future*, 6(5), 1–20. <https://doi.org/10.1029/2018EF000816>
- Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). *Increasing risk of compound flooding from storm surge and rainfall*. *Nature Climate Change*, 5, 1093–1097. <https://doi.org/10.1038/nclimate2736>



Ward, R. C., & Robinson, M. (2000). *Principles of hydrology* (4th ed.). McGraw-Hill.