ANALISIS KOMPARATIF MODEL RANDOM FOREST, LSTM, DAN HYBRID RF-LSTM DENGAN DATA AUGMENTED & FEATURES ENGINEERING UNTUK IMBALANCE DATA PADA SISTEM PERINGATAN DINI BENCANA BANJIR

Nico Flassy, Adza Zarif Nur Iskandar, Farhan Ariyanto, Aji Seto Arifianto
Teknik Informatika, Politeknik Negeri Jember
Jl. Mastrip, Krajan Timur, Sumbersari, Kec.Sumbersari,
Kabupaten Jember, 68121 Jawa Timur, Indonesia
nikoflassy@gmail.com

ABSTRAK

Fenomena bencana banjir yang terus meningkat memerlukan sistem peringatan dini yang presisi sebagai upaya preventif untuk mengurangi dampak yang ditimbulkan. Dalam pengembangannya, prediksi banjir berbasis *machine learning* menghadapi tantangan utama berupa karakteristik data meteorologi yang bersifat non-linear, memiliki ketergantungan temporal, serta distribusi data yang tidak seimbang. Penelitian ini bertujuan untuk mengevaluasi dan membandingkan kinerja tiga model, yaitu *Random Forest* (RF), *Long Short Term Memory* (LSTM), dan model *Hybrid* RF-LSTM dalam memprediksi kejadian banjir. Data yang digunakan berasal dari data BMKG Kabupaten Pasuruan yang diproses melalui *data augmentation* untuk mengatasi ketidakseimbangan kelas, dengan perbaikan rasio dari 1:11 menjadi 1:1,86 tanpa harus *undersampling*, serta didukung oleh proses *feature engineering* untuk memperkaya representasi data. Hasil eksperimen menunjukkan bahwa model *Hybrid* RF-LSTM memperoleh performa terbaik dengan nilai *Accuracy* sebesar 0,9667, *Precision* 1,0000, *Recall* 0,8333, F1-*Score* 0,9091, dan ROC-AUC 0,9792, yang menunjukkan kemampuannya dalam menangani dataset terbatas secara efektif. Sementara itu, model LSTM menunjukkan performa yang lebih rendah dengan *Accuracy* sebesar 0,7000 dan ROC-AUC 0,6696, yang mengindikasikan bahwa pendekatan *deep learning* cenderung kurang optimal pada kondisi dataset yang terbatas. Oleh karena itu, penelitian ini merekomendasikan penggunaan model *Hybrid* RF-LSTM atau *Random Forest* dalam sistem peringatan dini banjir pada kondisi data yang terbatas dan tidak seimbang.

Kata kunci : *Prediksi Banjir, Random Forest, LSTM, Hybrid RF-LSTM, Feature Engineering, Data Augmentation*

1. PENDAHULUAN

Fenomena bencana alam menjadi ancaman serius bagi keberlangsungan hidup makhluk hidup. Tidak hanya itu, fenomena ini juga mengganggu stabilitas ekonomi. Indonesia sebagai negara kepulauan yang terletak pada *Ring of Fire* dan memiliki iklim tropis menjadikannya sebagai negara yang cukup banyak terjadi bencana alam. Berdasarkan data dari Badan Penanggulan Bencana Daerah atau BPBD, pada tahun 2023 telah terjadi 5.400 kejadian bencana di Indonesia yang mana sangat jauh apabila dibandingkan dengan data bencana yang terjadi pada tahun 2014, yaitu sekitar 1.961 bencana [1].

Peningkatan tren bencana ini terjadi karena faktor perubahan cuaca & iklim setiap tahunnya atau yang dikenal sebagai bencana meteorologi [2].

Dalam konteks prediksi banjir, pendekatan berbasis data menjadi hal utama dalam mengembangkan sistem peringatan dini yang presisi. Data yang digunakan dalam sistem prediksi banjir pada umumnya berasal pada varibel meteorologi seperti curah hujan, suhu, kelembapan, kecepatan angin, dan parameter lain yang saling berinteraksi membentuk pola tertentu. Pola-pola ini sering kali dimanfaatkan untuk mengidentifikasi potensi terjadinya banjir. Namun demikian, data meteorologi memiliki karakteristik yang kompleks seperti memiliki sifat non-linear, memiliki ketergantungan temporal, dan menunjukan sifat fluktuasi yang tinggi antar waktu.

Sifat fluktuasi yang tinggi tersebut menyebabkan ketidakstabilan distribusi data. Dalam praktiknya, kejadian banjir (*flood*) merupakan peristiwa yang jarang terjadi dibandingkan kondisi normal (*non flood*). Hal tersebut mengakibatkan dataset memiliki ketimpangan distribusi antara data yang *flood* dan *non flood* meskipun data yang diambil sudah dari salah satu kabupaten dengan kejadian banjir tertinggi, yaitu Kabupaten Pasuruan. Selain itu, variasi nilai antar fitur yang berubah secara dinamis dalam periode tertentu menyebabkan pola kejadian banjir menjadi semakin sulit dikenali, terutama ketika jumlah sample terbatas.

Ketidakseimbangan distribusi ini kemudian menjadi tantangan dalam pengembangan model machine learning. Ketidakseimbangan tersebut menyebabkan model cenderung lebih banyak belajar dari data kelas mayoritas (*non flood*). Hal tersebut berpotensi menimbulkan kesalahan berupa false negative, yaitu kondisi ketika banjir tidak terdeteksi oleh model [3].

Kondisi ini menjadi hal yang krusial karena dapat berdampak buruk pada hasil prediksi model.

Untuk mengatasi masalah tersebut, diperlukan pendekatan yang mampu meningkatkan representasi data pada kelas minoritas tanpa menghilangkan karakteristik aslinya. Salah satu metode yang dapat

digunakan adalah *data augmentation*, yaitu teknik untuk menghasilkan variasi data abru dari data yang sudah ada. Pendekatan ini dapat dilakukan pada data *time series* dengan menambahkan *noise*, *scalling*, maupun pergeseran temporal untuk mensimulasikan variasi kondisi nyata sehingga distribusi data antara kelas *flood* dan *non flood* menjadi lebih [4].

Meskipun demikian, peningkatan kualitas dan distribusi data saja belum cukup untuk menghasilkan prediksi yang optimal. Data meteorologi yang kompleks, non-linear, dan memiliki ketergantungan temporal membutuhkan representasi fitur yang lebih informatif agar dapat dipahami dengan baik oleh model. Oleh karena itu, selain pemilihan algoritma, diperlukan proses *feature engineering* untuk mengekstraksi dan memperkaya informasi dari data mentah sehingga pola yang relevan dapat direpresentasikan dengan lebih jelas. Representasi fitur yang baik kemudian perlu didukung oleh model yang mampu memanfaatkan informasi tersebut secara optimal. Dalam hal ini, *Random Forest* dikenal efektif dalam menangani data tabular berdimensi tinggi serta memiliki ketahanan terhadap outlier, namun memiliki keterbatasan dalam menangani dinamika temporal data pada deret waktu. Di sisi lain, *Long Short Term Memory* (LSTM) dirancang untuk memproses data sekuensial dan mampu menangkap pola jangka panjang. Namun LSTM memiliki kelemahan yaitu memerlukan data yang besar sehingga akan cenderung kurang optimal pada dataset yang terbatas [5].

2. TINJAUAN PUSTAKA.

2.1. Banjir dan Mitigasi Bencana

Banjir merupakan salah satu bencana alam yang paling sering terjadi di Indonesia dan menimbulkan dampak signifikan terhadap kehidupan masyarakat, baik dari segi ekonomi, sosial, maupun lingkungan. Oleh karena itu, diperlukan upaya mitigasi yang efektif untuk meminimalisir risiko yang ditimbulkan. Prediksi banjir menjadi salah satu komponen penting dalam sistem mitigasi preventif, terutama dengan memanfaatkan teknologi berbasis data. Hal ini sejalan dengan penelitian yang menyatakan bahwa banjir merupakan bencana yang sering terjadi dan membutuhkan sistem prediksi yang akurat sebagai dasar mitigasi [6].

2.2. Prediksi Banjir Berbasis *Machine Learning*

Perkembangan teknologi *machine learning* telah memberikan kontribusi besar dalam meningkatkan akurasi prediksi bencana, termasuk banjir. Model *machine learning* mampu mengolah data dalam jumlah besar serta mengenali pola kompleks yang sulit dianalisis dengan metode konvensional. Penggunaan algoritma *machine learning* terbukti dapat meningkatkan akurasi prediksi dibandingkan metode tradisional karena kemampuannya dalam menangani data non-linear dan multidimensi [7].

Dalam implementasinya, model prediksi banjir biasanya menggunakan variabel seperti curah hujan, debit air, dan kondisi lingkungan sebagai parameter utama dalam menentukan potensi terjadinya banjir.

2.3. Algoritma *Random Forest* dalam Prediksi Banjir

Random Forest merupakan algoritma berbasis *ensemble* yang banyak digunakan dalam prediksi banjir karena kemampuannya dalam meningkatkan akurasi dan stabilitas model. Algoritma ini bekerja dengan menggabungkan banyak *decision tree* sehingga mampu mengurangi *overfitting*. Hasil penelitian menunjukkan bahwa *Random Forest* memiliki performa yang stabil dan lebih unggul dibandingkan metode lain seperti *Support Vector Machine* dalam prediksi banjir [8].

2.4. Algoritma *Long Short-Term Memory* (LSTM)

Long Short-Term Memory (LSTM) merupakan salah satu model *deep learning* yang dirancang untuk memproses data sekuensial. Model ini memiliki kemampuan dalam menangkap pola jangka panjang pada data *time-series* seperti curah hujan dan tinggi muka air. Penelitian menunjukkan bahwa sistem prediksi berbasis LSTM mampu memberikan hasil yang cepat dan akurat dalam memprediksi banjir sehingga dapat membantu mengurangi dampak negatif yang ditimbulkan [9].

2.5. Model *Hybrid Random Forest* dan LSTM

Pendekatan *hybrid* antara *Random Forest* dan LSTM dikembangkan untuk mengatasi keterbatasan masing-masing model. Model ini menggabungkan kemampuan *Random Forest* dalam menangani data non-linear dengan kemampuan LSTM dalam memahami pola temporal. Penelitian menunjukkan bahwa pendekatan *hybrid* ini mampu meningkatkan akurasi prediksi banjir dengan memanfaatkan data geospasial dan temporal secara bersamaan [10].

2.6. *Class Imbalance* dan *Data Augmentation*

Ketidakimbangan *class* pada data yang digunakan untuk klasifikasi banjir merupakan tantangan fundamental utama. Jumlah data pada *class non-flood* memiliki perbandingan jumlah yang sangat jauh dengan data pada *class flood.* Kondisi tersebut menyebabkan model menjadi bias terhadap class dominan sehingga dapat menyebabkan performa menurun untuk mendeteksi kejadian banjir sebenarnya [11].

Beberapa cara untuk mengatasi permasalahan tersebut antara lain adalah dengan menggunakan pendekatan *data augmentation*. Teknik ini dapat meningkatkan jumlah sampel *flood* dari 40 menjadi 240 (6x lipat) tanpa mengurangi data asli.

2.7. *Temporal Dependency* dan *Time-Based* Sampling

Data meteorologi temporal yang digunakan dalam prediksi banjir memiliki karakteristik *time series* dengan dependensi temporal. Hal tersebut

berarti bahwa nilai pada suatu periode waktu dipengaruhi oleh nilai-nilai sebelumnya. Berbeda dengan badan tabular konvensional yang pengolahan data deret waktunya memerlukan strategi pembagian *dataset* yang khusus untuk mencegah *data leakage* [12].

Penggunaan teknik *random sampling* menyebabkan informasi masa depan bocor ke dalam proses pelatihan sehingga akan berakibat pada overestimasi performa model. Pemahaman terhadap ketergantungan temporal telah dibahas oleh Nurcahyo dan Redjeki yang menggunakan pendekatan *esemble learning* berbasis indikator atmosfer dan menunjukkan bahwa pola temporal pada data *time series* mampu ditangkap dengan baik menggunakan *feature engineering* yang tepat [13].

Oleh karena itu, pemodelan yang paling tepat adalah dengan menggunakan teknik *engineering* yang baik.

2.8. Evaluasi Model untuk *Imbalanced Dataset*

Tolak ukur pada performa pada *dataset* ini menggunakan metrik *accuracy* yang mana hal tersebut dapat menimbulkan interpretasi yang menyesatkan atau *misleading*. Hal tersebut terjadi karena model dapat dapat mencapai akurasi tinggi dengan selalu memprediksi kelas mayoritas [14].

Oleh karena itu, diperlukan metrik yang lebih representatif seperti *F1-Score* yang merupakan rata-rata dari nilai *precision, recall dan Area Under the ROC Curve* (AUC). Penerapan model pada peringatan dini bencana membuat *recall* (sensitivitas) merupakan metrik yang krusial untuk meminimalkan *negative false* (kegagalan dalam melakukan deteksi yang dapat berpotensi fatal.

2.9. *Feature Engineering* dan Karakteristik Fitur

Penelitian ini menggunakan total 25 fitur yang terdiri dari 6 fitur original dan 19 fitur hasil *engineering*. Fitur original meliputi temperatur (t), curah hujan (rr), kelembapan (hu), kecepatan angin (ws), arah angin (wd), dan lama penyinaran matahari (ss). Fitur hasil *engineering* mencakup fitur temporal *(month, day_of_year, is_weekend*), fitur *rainfall* kumulatif (*rain_3h, rain_6h, rain_12h, rain_24h, rain_48h*), fitur *alert* (*temp_alert, humidity_alert, visibility_alert, rain_intensity*), fitur interaksi (*temp_humidity, wind_vis, rain_humidity, rain_temp*), serta fitur lag (*hu_lag1, ws_lag1, rain_3h_lag1*).

Hasil menunjukkan bahwa *feature engineering* yang kompleks sangat membantu model *Random Forest* mencapai ROC-AUC 0,9733 dan precision 1,000. Hal ini mengindikasikan bahwa fitur meteorologi permukaan dengan *features engineering* yang tepat mampu memberikan konteks yang kuat untuk membedakan kondisi *flood* dan *non-flood* sehingga hal ini bisa dimanfaatkan oleh model *tree-based* untuk melakukan prediksi yang akurat.

2.10. *Data Augmentation* untuk *Time Series*

Berbeda dengan SMOTE yang menciptakan data sintetis dengan cara mencampur dua data asli, *data augmentation* untuk *time series* menciptakan data baru yang independen dengan cara mengubah data asli satu per satu Teknik yang paling umum digunakan meliputi *gaussian noise, magnitude scaling, dan time warping* untuk mensimulasikan variasi *measurement error*, intensitas fenomena, dan pergeseran temporal [15].

Data augmentasi dengan variasi kurang lebih 2-3% pada data *time series* dapat meningkatkan generalisasi model [15].

Sementara itu, Iwana dan Uchida (2021) menjelaskan bahwa ketiga pendekatan tersebut bersifat komplementer dalam menciptakan variasi dataset dengan mempertahankan stuktur temporal asli [16]

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen (*Experimental Research*) yang bertujuan untuk menganalisis dan membandingkan kinerja beberapa model *machine learning*, yaitu *Long Short-Term Memory* (LSTM), *Random Forest* (RF), serta model *Hybrid* RF-LSTM dalam memprediksi potensi banjir di Kabupaten Pasuruan.

Pendekatan ini dipilih karena mampu memberikan hasil yang objektif melalui pengukuran kinerja model berdasarkan metrik evaluasi tertentu. Proses penelitian dilakukan secara sistematis mulai dari tahap pengumpulan data, pra-pemrosesan data, implementasi model, hingga evaluasi dan validasi model.

3.1. Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data multivariat yang mencakup variabel klimatologis dan geografis pada wilayah kabupaten Pasuruan. Data terbagi menjadi 2 yaitu data klimatologi dan data geografis. Data klimatologi meliputi curah hujan, suhu, kelembapan, dan kecepatan angin yang diperoleh dari data BMKG Kabupaten Pasuruan.

3.2. Pra-pemrosesan Data (*Pre-processing*)

Tahap pra-pemrosesan data bertujuan untuk menyiapkan data agar dapat digunakan secara optimal dalam proses pelatihan model. Adapun tahapan yang dilakukan meliputi:

a. Pembersihan data (*data cleaning*) dari nilai hilang (*missing value*) dan data anomali yang mengandung nilai 8888 dan 9999. Nilai tersebut merepresentasikan *missing value*. Nilai ini akan diganti dengan media masing masing fitur
b. *Data augmentation* untuk kelas minoritas karena data yang digunakan memiliki rasio ekstrem yaitu 1:11. Teknik yang digunakan adalah *gaussian noise* dengan variasi kurang lebih 2%,

magnitude scaling 0.95-1.05 pada seluruh fitur, dan *data offset* kurang lebih 2 hari untuk variasi data temporal.

c. *Feature engineering* dilakukan untuk mengekstraksi informasi tambahan dari data mentah meliputi fitur temporal (*month, is_wet_season*), fitur *rainfall* (*rr_cum_3d, rr_cum_7d, rr_max_3d, rr_intensity*), fitur *alert* (*alert_heavy, alert_3d_pattern*), fitur interaksi (*temp_humidity, rainfall_wind*), serta *lag features* (*hu_lag1, ws_lag1, rain_3h_lag1)*
d. Normalisasi atau transformasi normalisasi *min max scaler*. Untuk LSTM, data diubah ke format *sequence* (*timesteps* = 10)
e. Pembagian *dataset* menjadi data latih (*training set*) sebesar 70%, data validasi (*validation set*) sebesar 15%, dan data uji (*testing set*) sebesar 15%

3.3. Model LSTM

Model LSTM digunakan sebagai komponen yang menangani data *time series* atau data deret waktu. Arsitektur pada model ini terdiri dari beberapa layer yang antara lain input layar, beberapa *hidden layer*, hingga *dense layer*. LSTM bertugas membaca pola dari fluktuasi curah hujan dari waktu ke waktu.

3.4. Model *Random Forest*

Model *random forest* diimplementasikan sebagai *ensemble trees*. Model ini bekerja dengan membangun *decision trees* selama masa training dan menghasilkan output berdasarkan rata rata prediksi (regresi) atau suara paling banyak (klasifikasi).

3.5. Model *Hybrid*

Model *Hybrid* merupakan skema *ensemble* dengan menggabungkan prediksi probabilitas dari dua model yang dilatihan secara terpisah, yaitu *Random Forest* dan LSTM. *Random Forest* menghasilkan probabilitas prediksi berdasarkan fitur statis sedangkan LSTM menghasilkan probabilitas prediksi berdasarkan *sequence* temporal. Kedua probabilitas tersebut akan digabungkan menggunakan *weighted averaging* dengan bobot 0.5 untuk RF dan 0.5 untuk LSTM. Hasil *ensemble* tersebut kemudian diklasifikasikan menggunakan *threshold* 0.45 untuk menghasilkan prediksi akhir. Dengan melakukan pendekatan ini model akan memanfaatkan kekuatan RF dalam menangani fitur terstruktur dan kekuatan LSTM dalam menangkap pola temporal tanpa memerlukan *fully connected layer* tambahan.

3.6. Evaluasi Model

Kinerja model diukur menggunakan metrik standar untuk memastikan presisi prediksi. Pada penelitian ini efektivitas algoritma dilakukan menggunakan beberapa matrik evaluasi yaitu:

a. *Accuracy* digunakan untuk mengukur proporsi prediksi benar dari total data. Metrik ini memberikan gambaran umum mengenai kemampuan model dalam mengklasifikasikan kondisi *flood* dan *non-flood* secara tepat.
b. *Precision* digunakan untuk mengukur proporsi prediksi *flood* yang benar dari total prediksi *flood*. Metrik ini penting untuk meminimalkan false alarm yaitu melakukan prediksi banjir namun pada faktanya tidak terjadi.
c. *Recall (Sensitivity*) digunakan untuk mengukur kemampuan model mendeteksi kejadian flood yang sebenarnya.
d. *F1-Score* digunakan untuk menghitung rata-rata harmonik antara *precision* dan *recall*. Metrik ini memberikan gambaran keseimbangan antara kemampuan model dalam mengurangi *false alarm* dan kemampuan dalam mendeteksi seluruh kejadian banjir.
e. ROC-AUC digunakan untuk mengukur kemampuan model dalam membedakan kelas *flood* dan *non-flood*. Semakin tinggi nilainya dengan mendekati nilai 1 maka menunjukkan bahwa kemampuan diskriminasi model tersebut yang sangat baik.

3.7. Validasi

Penelitian ini menerapkan teknik hold-out *validation* dengan pembagian data menjadi 3 bagian yaitu *training set* dengan persentase 70%, *validation set* persentase 15%, dan *testing set* dengan persentase 15%. *Training set* digunakan untuk melatih model, *validation set* digunakan untuk *tuning hyperparameter* dan *early stopping*, serta *testing set* digunakan untuk evaluasi akhir model selama proses *training*.

4. HASIL DAN PEMBAHASAN

4.1. Karakteristik *Dataset*

Dataset yang digunakan pada penelitian ini diambil dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) Kabupaten Pasuruan yang terdiri dari 486 data dengan periode data Bulan Maret 2023 sampai Maret 2024. Fitur *dataset* yang digunakan antara lain *temperature* (t), curah hujan (rr), kelembaban udara (hu), kecepatan angin (ws), arah angin (wd), lama penyinaran matahari (ss), & *flood* (f). Semua parameter tersebut diambil dari data online BMKG kecuali *flood* (f) yang diambil dari validasi *scrapping* berita di internet. Validasi untuk *flood* diambil dengan melewati beberapa proses ketat yaitu mulai dari *scraping* hingga validasi dengan *threshold* 3 berita untuk memastikan bahwa berita tersebut valid dari beberapa sumber.

Setelah proses *feature engineering*, jumlah fitur bertambah menjadi 25 fitur antara lain terdiri dari fitur temporal, fitur *rainfall* kumulatif & intensitas, fitur *alert*, fitur interaksi antar variabel, serta *lag features* untuk menangkap dependensi temporal.

4.2. Karakteristik *Class*

Data menunjukkan bahwa kondisi *imbalance* ekstrem dengan rasio 1:11 antara *class flood* dan *non flood*. Dari total total 486 data, ada 446 data yang *non-*

flood dan hanya 40 data *flood*. Kondisi ini cukup menjadi tantangan dikarenakan dapat menyebabkan bias terhadap kelas data mayoritas yang mengakibatkan sulitnya mendeteksi kelas data minoritas. Proses augmentasi yang dilakukan pada penelitian ini membuat data flood menjadi meningkat menjadi 240 data sehingga menghasilkan total 686 data dengan rasio 1:1.86.

Tabel 1. Karakteristik distribusi kelas dataset sebelum dan sesudah data *augmented*

Kondisi	*Class Flood*	*Class Non-Flood*	Rasio
Sebelum *Augmented*	40	446	1:11
Setelah *Augmented*	240	446	1: 1.86

4.3. Strategi *Data Augmentation*

Permasalahan *class* yang *imbalance* dapat diatasi dengan beberapa cara, salah satunya dengan strategi *data augmentation*. Strategi ini bertujuan untuk menambahkan data minoritas (*flood*) supaya memiliki perbandingan yang tidak terlalu jauh dengan kelas mayoritas saat ini (*non flood*). Teknik augmentasi ini menggunakan 3 pendekatan yaitu *gaussian noise*, *magnitude scaling*, dan *data offset.*

4.4. *Time-Based Data Splitting*

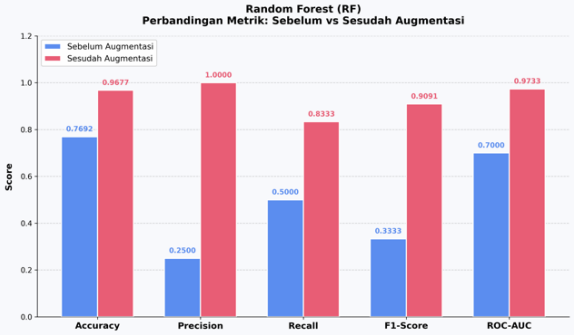
Data dari BMKG tersebut merupakan data dengan sifat *time series* sehingga pembagian *dataset* dilakukan secara kronologis (*time base split*) dengan proporsi 70% untuk data *training*, 15% untuk data validasi, dan 15% untuk data *testing*. Pendekatan *splitting* ini dapat digunakan untuk mencegah *data leakage* dan mensimulasikan kondisi prediksi *real-time* dimana model dilatih untuk memprediksi kejadian di masa depan.

4.5. Normalisasi Data

Data yang ada di beberapa fitur tersebut memiliki *value* yang berbeda beda. Untuk menyeragamkan *value* tersebut ke rentang yang sama, yaitu 0 sampai 1 menggunakan *min max scaling* maka diperlukannya normalisasi data. Beberapa fitur yang dinormalisasi adalah t, hu, ws, dan ss. Sedangkan fitur yang tidak dinormalisasi adalah *date*, wd, dan f yang mana f sudah berbentuk biner.

4.6. *Random Forest* (RF)

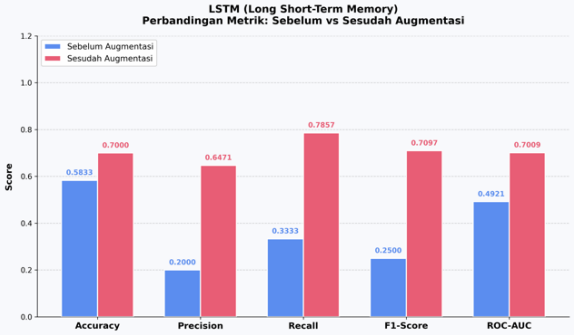
Model *Random Forest* dengan 500 estimators dan max depth 20 menghasilkan performa sangat baik dengan *accuracy* 96,77%, *precision* 100%, *recall* 83,33% dan *F1-Score* 0,9091.



Gambar 1. Perbandingan metrik model *random forest* sebelum dan sesudah augmentasi.

4.7. *Long Short-Term Memory* (LSTM)

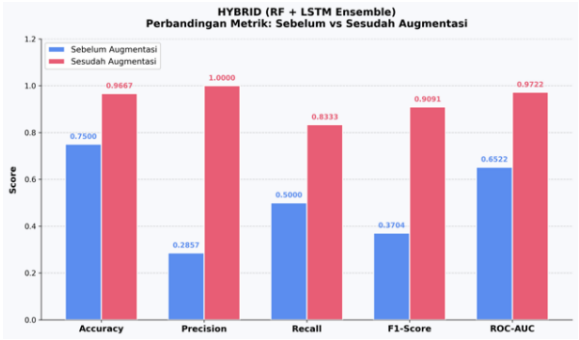
Arsitektur LSTM dengan 2 layer (64 dan 32 unit), *dropout* 0.3, dan *sequence* 10 hari menunjukkan bahwa performa terbatas dengan *accuracy* 70%, *precision 64,71%, dan F1-Score* 0,7097. Nilai ROC-AUC pada model ini adalah 0,6696 yang mengindikasikan kemampuan diskriminasi yang lemah dibandingkan dengan model yang lain. Model tersebut mengalami *underfitting* akibat dataset yang terlalu kecil sehingga model kesulitan dalam belajar melalui arsitektur *deep learning* yang dimiliki.



Gambar 2. Perbandingan metrik model LSTM sebelum dan sesudah augmentasi.

4.8. *Hybrid* RF-LSTM

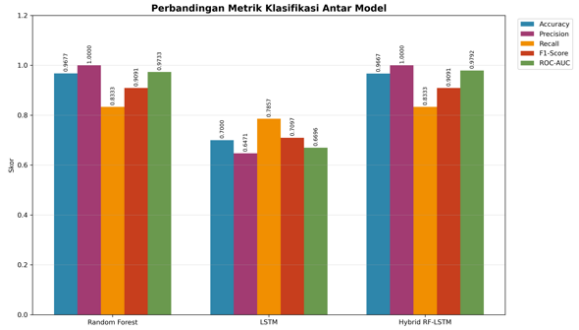
Model *Hybrid* ini dengan *weighted averaging* menghasilkan *accuracy* 96,67%, *recall* yang sempurna yaitu 100%, recall 83,33%, dan F1-*Score* 0,9091. Nilai ROC-AUC 0,9792 merupakan nilai tertinggi dibandingkan *Random Forest* & LSTM yang menandakan bahwa kombinasi kedua model tersebut berhasil meningkatkan kemampuan diskriminasi. *Precision* sempurna diwarisi oleh model *tree-based* yaitu *Random Forest* yang dilengkapi dengan kontribusi LSTM yang memberikan peningkatan ROC-AUC dari 0,9733 pada RF menjadi 0,9792 di model *hybrid* ini.



Gambar 3. Perbandingan metrik model *Hybrid* sebelum dan sesudah augmentasi

4.9. Dominasi *Random Forest* dan *Hybrid* RF-LSTM pada Metrik Kunci

Berdasarkan analisa dari 5 metrik evaluasi utama, *Hybrid* RF-LSTM dan *Random Forest* mendominasi dengan ROC-AUC masing-masing dengan nilai 0,9792 dan 0,9733. Kedua model tersebut mencapai *precision* sempurnasdengan angka 1,000 yang menunjukkan bahwa tidak ada *false alarm* dan F1-*Score* 0,9091 yang mengindikasikan keseimbangan sempurna.



Gambar 4. Perbandingan metrik klasifikasi antar model

Hal ini menunjukkan bahwa model *tree-based* dengan *feature engineering* yang sangat cocok untuk dataset terbatas. Sebaliknya, LSTM dengan ROC-AUC 0,6696 yang menunjukkan performa yang tidak baik. Hal ini berarti bahwa ketidakmampuan *deep learning* pada dataset kecil.

4.10. Keterbatasan LSTM pada Dataset Kecilv

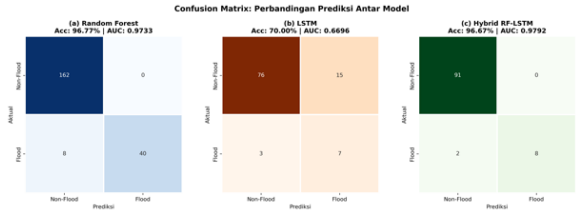
LSTM menunjukkan accuracy 0,7000 dan ROC-AUC 0,6696 yang mengindikasikan kesulitan dalam belajar akibat dataset yang terlalu kecil (*underfitting*). Dataset yang minim tersebut yang mengakibatkan kurang unggulnya performa pada model LSTM.

4.11. Keunggulan Model *Hybrid*

Hybrid RF-LSTM unggul dengan nilai ROC-AUC 0,9792 sebagai nilai tertinggi. Model *ensemble* dengan *weighted averaging* 0,5 berhasil menggabungkan probabilitas RF yang kuat dengan insight temporal LSTM melampaui performa kedua *single* model tersebut.

4.12. Analisis *Confusion Matrix*

Visualisasi pada *confusion matrix* menunjukkan pola yang signifikan antar ketiga model. *Random Forest* menunjukkan performa sangat baik dengan 162 *true negative* dan 40 *true positive*, tanpa *false positive*, dan 8 *false negative*. LSTM menunjukkan banyak *false positive* yaitu 15 yang mengindikasikan *false alarm* tinggi dengan 76 *true negative* dan 7 *true positive*. *Hybrid* RF-LSTM menunjukkan performa optimal dengan 91 *true negative*, 8 *true positive*, tanpa *false positive*, dan hanya 2 *false negative*. *Random Forest* dan *Hybrid* memiliki *precision* sempurna, sementara LSTM menghasilkan tingkat *false positive* yang tinggi akibat *underfitting* pada dataset kecil.



Gambar 5. *Confusion matrix* perbandingan model *Random Forest*, LSTM, dan *Hybrid* RF-LSTM

4.13. Karakteristik Optimal *Random Forest*

Random Forest dengan nilai *Recall* 0,8333 dan precision 1,000 menunjukkan bahwa model ini sangat ideal untuk digunakan dalam sistem peringatan dini bencana. *Precision* yang sempurna mengindikasikan tidak ada *false alarm*, sementara *recall* dengan nilai 83,33% menunjukkan bahwa hanya ada 16,67% *missed detection*. 25 fitur terstruktur pada *feature engineering* kompleks ini berhasil dimanfaatkan dengan baik oleh model ini sehingga dapat menghasilkan performa yang tinggi.

4.14. Karakteristik Terbatas LSTM

LSTM dengan *precision* 64,71% menunjukkan tingkat *false alarm* yang tinggi. Rendahnya *precision* ini bukan merupakan indikasi sensitivitas tinggi, melainkan indikasi bahwa terjadinya *underfitting* yang disebabkan oleh dataset yang kecil untuk memenuhi kebutuhan arsitektur *deep learning*. Hasil ini menunjukkan bahwa LSTM tidak cocok untuk dataset terbatas meskipun dirancang khusus untuk data *time series*.

4.15. Karakteristik Superior *Hybrid* RF-LSTM

Model *Hybrid* mencapai *recall* 83,33% dengan *precision* 100% menunjukkan bahwa model ini mendominasi dengan kontribusi RF yang kuat dalam *ensemble*. *Weighted averaging* 0,5:0,5 berhasil menggabungkan probabilitas *Random Forest* yang kuat dengan *insight* temporal LSTM dengan hasil ROC-AUC 0,9792. Nilai tersebut meningkat dibandingkan dengan nilai ROC-AUC pada RF murni yang bernilai 0,9733. *Threshold* 0,45 yang lebih rendah dari *default* (0,5) memungkinkan model sedikit lebih sensitif terhadap sinyal positif dari model LSTM

tanpa mengkorbankan *precision*. Karakteristik ini menempatkan model Hybrid sebagai model terbaik dengan keseimbangan sempurna dan deteksi banjir yang tinggi.

4.16. Karakteristik Model

Hasil eksperimen ini menunjukkan perbedaan fundamental dalam kemampuan model pada dataset terbatas. *Random Forest* berhasil mencapai ROC-AUC 0,9733 dan *Precision* 1,000 mengindikasikan kekuatan model *tree-based* dengan *feature engineering* yang tepat pada data kecil. Sebagai model *ensemble* yang bekerja pada fitur flat, model RF mampu memanfaatkan 25 fitur secara efektif tanpa memerlukan pola temporal eksplisit dengan bantuan *feature engineering* yang kompleks. Berbeda dengan *Random Forest*, LSTM menunjukkan performa terbatas dengan ROC-AUC 0,6696 yang mengindikasikan terjadi *underfitting* akibat dataset yang kecil. Kemampuan LSTM dalam memproses *sequence* 10 hari tidak berhasil mengidentifikasikan pola yang bermakna karena *deep learning* pada umumnya memerlukan data yang cukup banyak untuk mendapatkan hasil yang maksimal. Model *Hybrid* RF-LSTM mencapai performa tertinggi dengan ROC-AUC 0,9792. Model *ensemble* dengan *weighted averaging* 0,5:0,5 berhasil menggabungkan kekuatan RF dalam menangani fitur statis dengan *insight* temporal LSTM melampui performa kedua model individual. Peningkatan ROC-AUC dari 0,9733 pada RF menjadi 0,9792 pada *Hybrid* menunjukkan kontribusi LSTM memberikan tambahan diskriminasi yang berharga dalam kombinasi *ensemble*.

4.17. Keterbatasan

Penelitian ini memiliki beberapa keterbatasan yang dapat mempengaruhi hasil model. Pertama, jumlah *dataset* yang terbatas yaitu 687 setelah augmentasi. Hal tersebut menyebabkan model LSTM tidak dapat belajar secara optimal dan mengalami *underfitting* karena arsitektur *deep learning* yang memerlukan jumlah dataset yang banyak. Namun, model *Random Forest* dan *Hybrid* berhasil mengatasi keterbatasan ini melalui pendekatan *feature engineering* yang kuat dan data augmentasi.

Kedua, fitur cuaca yang digunakan seperti suhu (t), curah hujan (rr), kelembapan (hu), kecepatan angin (ws), dan arah angin (wd) dan lama penyinaran matahari (ss) merupakan parameter permukaan. Meskipun *feature engineering* berhasil meningkatkan performa RF dan *Hybrid*, integrasi fitur geospasial seperti elevasi topografi, indeks vegetasi, tutupan lahan, dan parameter hidrologi seperti debit sungai masih direkomendasikan untuk peningkatan pada implementasi selanjutnya.

Ketiga, karakteristik banjir pada dataset terjadi tiba-tiba dalam kurun waktu yang singkat sehingga membuat pola temporal yang panjang kurang relevan sehingga menjadi salah satu faktor ketidakmampuan LSTM meskipun dirancang untuk *time series*.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian komparatif antara model Random Forest, LSTM, dan Hybrid RF-LSTM untuk prediksi banjir di Kabupaten Pasuruan, dapat disimpulkan bahwa model Hybrid RF-LSTM menunjukkan performa terbaik dengan ROC-AUC 0,9792, Precision 1,000, dan F1-Score 0,9091. Random Forest menunjukkan performa yang sangat kompetitif dengan ROC-AUC 0,9733 dan Precision yang sama dengan model Hybrid yaitu 1,000 yang mengindikasikan bahwa model tree-based pada dataset terbatas dengan feature engineering yang tepat. Berbeda dengan kedua model tersebut, LSTM menunjukkan performa yang terbatas dengan ROC-AUC 0,6696 yang mengindikasikan ketidakmampuan model deep learning pada dataset yang kecil. Hal ini menunjukkan bahwa fitur meteorologi dengan feature engineering kompleks mampu membedakan kejadian banjir dengan sangat baik tanpa memerlukan dependensi temporal panjang. Selain itu, penerapan data augmented dalam penelitian ini terbukti mampu mengurangi ketidakseimbangan distribusi data yang dapat meningkatkan representasi kelas minoritas (flood). Oleh karena itu, disarankan menggunakan model Hybrid RF-LSTM atau Random Forest untuk sistem peringatan dini banjir dengan kondisi dataset terbatas dengan melakukan penambahan fitur menggunakan pendekatan feature engineering dan penambahan beberapa fitur geospasial seperti elevasi topografi, indeks vegetasi, dan parameter hidrologi untuk peningkatan performa pada implementasi berikutnya.

DAFTAR PUSTAKA

[1] Badan Nasional Penanggulangan Bencana, “Data Bencana Indonesia 2023,” 2024. Accessed: Mar. 30, 2026. [Online]. Available: https://bpbd.muarojambikab.go.id/public/deploy/pdf/1740371991_eb771516c83e2df0e0c5.pdf

[2] Badan Riset dan Inovasi Nasional (BRIN), “Badan Riset dan Inovasi Nasional (BRIN).” Accessed: Mar. 30, 2026. [Online]. Available: https://www.brin.go.id/news/112132/cuaca-ekstrem-meningkat-di-indonesia-salah-satu-indikasi-perubahan-iklim

[3] R. Muslim Sinaga, A. Afrrahman S. Effendi, N. Latifah Hasibuan, F. Asro Harahap, and F. Ramadhani, “EVALUASI PERBANDINGAN TEKNIK OVERSAMPLING TERHADAP KINERJA RANDOM FOREST DALAM MENDETEKSI TRANSAKSI HARI SPESIAL DI E-COMMERCE,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 9, no. 3, pp. 5317–5324, May 2025, doi: 10.36040/jati.v9i3.14156.

[4] K. Handayani, E. Erni, B. Lailiah, and R. Sa’adah, “KLASIFIKASI KANKER PAYUDARA MENGGUNAKAN EXTRA TREE DENGAN SMOTE,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 6,

pp. 3100–3105, Jan. 2024, doi: 10.36040/jati.v7i6.8797.

[5] A. Pambudi, D. Aryani, and R. N. Sunarto, "Rainfall Intensity Prediction Using LSTM and Random Forest Hybrid Model," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems*, vol. x, No.x, pp. 1–5, doi: 10.22146/ijccs.xxxx.

[6] T. Prasetyo, "Analisis Prediksi Risiko Banjir Menggunakan Algoritma Random Forest," *Jurnal Teknik Informatika Unika ST. Thomas (JTIUST)*, vol. 10, no. 02, pp. 2657–1501, 2025.

[7] I. Maulita, C. Raras, A. Widiawati, and A. Mu'amar Wahid, "Analisis Komparatif Linear Regression, Random Forest, dan Gradient Boosting untuk Prediksi Banjir," *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, vol. 4, no. 8, pp. 369–379, 2024, doi: 10.52436/1.jpti.599.

[8] I. A. Purnomo, J. Indra, E. E. Awal, and T. Rohana, "Analisis Prediksi Banjir di Indonesia Menggunakan Algoritma Support Vector Machine dan Random Forest," *Journal of Information System Research (JOSH)*, vol. 6, no. 1, pp. 219–228, Oct. 2024, doi: 10.47065/josh.v6i1.5958.

[9] B. D. Handika, I. Y. Rakhmawati, and F. I. Kurniawan, "Penerapan Algoritma Machine Learning Untuk Memprediksi Bencana Alam Banjir Menggunakan Algoritma SVM dan LSTM," *SUBMIT: Jurnal Ilmiah Teknologi Infomasi dan Sains*, vol. 5, no. 1, pp. 16–21, Jun. 2025, doi: 10.36815/submit.v5i1.3969.

[10] Z. F. Jailani and D. Nurmadewi, "HYBRID MACHINE LEARNING PREDIKSI BANJIR MENGGUNAKAN LSTM DAN RANDOM FORESTS PADA GEODATA HYBRID MACHINE LEARNING PREDICTS FLOODING USING LSTM AND RANDOM FORESTS ON GEODATA," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 8, no. 1, 2025.

[11] Haibo He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009, doi: 10.1109/TKDE.2008.239.

[12] M. Asif, M. M. Kuglitsch, I. Pelivan, and R. Albano, "Review and Intercomparison of Machine Learning Applications for Short-term Flood Forecasting," *Water Resources Management*, vol. 39, no. 5, pp. 1971–1991, Mar. 2025, doi: 10.1007/s11269-025-04093-x.

[13] A. Wilis Nurcahyo and S. Redjeki, "ANALISIS REGRESI NON-LINIER PRODUKSI PERIKANAN TANGKAP MENGGUNAKAN PENDEKATAN ENSEMBLE LEARNING BERBASIS INDIKATOR ATMOSFER," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 10, no. 2, pp. 2624–2629, Mar. 2026, doi: 10.36040/jati.v10i2.17610.

[14] H. Kaur, H. S. Pannu, and A. K. Malhi, "A Systematic Review on Imbalanced Data Challenges in Machine Learning," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–36, Jul. 2020, doi: 10.1145/3343440.

[15] T. T. Um *et al.*, "Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, New York, NY, USA: ACM, Nov. 2017, pp. 216–220. doi: 10.1145/3136755.3136817.

[16] B. K. Iwana and S. Uchida, "An empirical survey of data augmentation for time series classification with neural networks," *PLoS One*, vol. 16, no. 7, p. e0254841, Jul. 2021, doi: 10.1371/journal.pone.0254841.