

Perbandingan Algoritma Machine Learning untuk Prediksi Multi-Output Variabel Osean-Atmosfer

Raisa Maulidia¹, Willdan Aprizal Arifin^{*2}, Herman Syafri³

^{1,2,3}Universitas Pendidikan Indonesia, Jl. Dr. Setiabudi No.229, Isola, Kec. Sukasari, Kota Bandung, Jawa Barat 40154

e-mail: ¹raisamld@upi.edu, ^{*2}willdanarifin@upi.edu, ³herman.syafri@upi.edu

Abstrak

Prediksi variabel osean-atmosfer merupakan komponen penting dalam mendukung keselamatan dan efisiensi aktivitas maritim. Kompleksitas data osean-atmosfer yang bersifat multivariat dan dinamis memerlukan pendekatan komputasional yang mampu menangkap hubungan non-linear dan temporal secara simultan. Penelitian ini membandingkan performa tiga algoritma machine learning, yaitu Long Short-Term Memory (LSTM), Random Forest, dan XGBoost dalam memprediksi multi-output variabel osean-atmosfer menggunakan data Automatic Weather Station (AWS) periode 2022–2025. Tahapan penelitian meliputi pra-pemrosesan, normalisasi menggunakan StandardScaler, pembagian data (90% latih dan 10% uji), pelatihan model teroptimasi, serta evaluasi menggunakan RMSE dan R². Hasil penelitian menunjukkan bahwa XGBoost memiliki performa terbaik pada sebagian besar variabel dengan nilai RMSE terendah pada windspeed (0,77), waterlevel (0,12), RH (2,40), dan winddir (28,79), serta nilai R² tertinggi masing-masing sebesar 0,840; 0,940; 0,870; dan 0,730. LSTM menunjukkan performa terbaik pada variabel watertemp dengan RMSE sebesar 0,31 dan R² sebesar 0,814. Sementara itu, Random Forest memiliki performa yang relatif lebih rendah dengan nilai R² berkisar antara 0,680 hingga 0,982 tergantung variabel. Secara keseluruhan, XGBoost terbukti paling konsisten dan efektif dalam menangani prediksi multi-output variabel osean-atmosfer yang kompleks dan non-linear.

Kata kunci—LSTM, machine learning, multi-output, Random Forest, XGBoost

1. PENDAHULUAN

Isu perubahan iklim global saat ini menjadi sorotan penting dalam permasalahan lingkungan karena dampaknya yang semakin meluas ke berbagai sektor kehidupan [1]. Dampak perubahan iklim juga terasa hingga level regional seperti di Indonesia sebagai negara kepulauan yang sebagian besar wilayahnya adalah laut [2]. Prediksi yang akurat terhadap parameter osean-atmosfer, seperti suhu udara, kelembapan, kecepatan angin, maupun tinggi gelombang, menjadi kebutuhan penting dalam mendukung keselamatan dan efisiensi kegiatan di laut. Analisis beberapa parameter secara bersamaan dapat memberikan gambaran yang lebih menyeluruh mengenai kondisi serta dinamika yang terjadi di laut [3]. Namun, data osean-atmosfer memiliki karakteristik yang kompleks karena dipengaruhi oleh banyak faktor dan dapat mengalami fluktuasi yang tinggi. Kompleksitas tersebut menimbulkan tantangan dalam proses pemodelan prediksi, sehingga diperlukan pendekatan komputasional yang tepat.

Seiring dengan perkembangan teknologi, penelitian di bidang prediksi atmosfer-oseanografi terus mengalami kemajuan. Teknologi penginderaan dan pemantauan cuaca memungkinkan ketersediaan data atmosfer-oseanografi yang lebih detail dan *real-time* melalui Automatic Weather Station (AWS). AWS menghasilkan data deret waktu dengan resolusi tinggi yang bersifat multivariat serta berkelanjutan [4]. Data ini sangat potensial untuk diolah

menggunakan algoritma *machine learning* agar menghasilkan prediksi yang akurat. Selain itu, pendekatan *multi-output prediction* juga semakin relevan karena mampu memprediksi beberapa variabel secara simultan. Dengan demikian, hasil prediksi yang diperoleh lebih komprehensif dan bermanfaat untuk mendukung pengambilan keputusan di sektor kelautan.

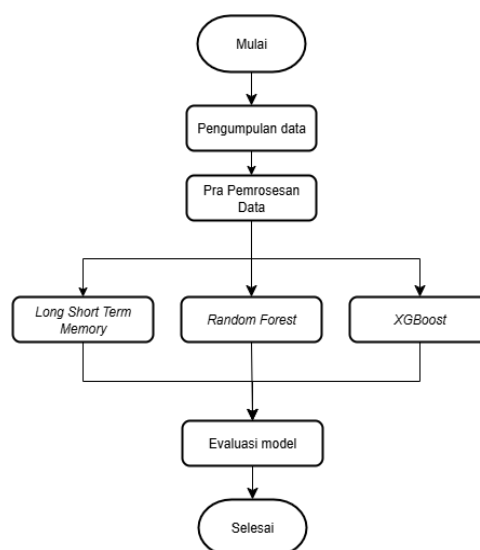
Berbagai algoritma *machine learning* telah digunakan untuk menangani permasalahan prediksi data kompleks, masing-masing dengan karakteristik dan keunggulannya. *Long Short-Term Memory* (LSTM) banyak diterapkan dalam pemodelan peramalan karena kemampuannya mempelajari pola temporal serta mempertahankan informasi jangka panjang secara efektif [5][6]. *Random Forest* dikenal sebagai algoritma berbasis *ensemble* yang handal dalam mengatasi *overfitting* dan bekerja baik pada data nonlinier [7]. Sementara itu, XGBoost merupakan algoritma *boosting* yang terkenal dengan efisiensi komputasi dan akurasi tinggi, terutama pada data dengan variabel yang kompleks [8]. Oleh karena itu, perbandingan ketiga algoritma ini penting untuk mengetahui performa terbaik dalam konteks prediksi *multi-output* variabel osean-atmosfer.

Meskipun masing-masing algoritma telah banyak diaplikasikan di berbagai bidang, penelitian yang berfokus pada perbandingan kinerja LSTM, *Random Forest*, dan XGBoost untuk prediksi *multi-output* variabel osean-atmosfer menggunakan data AWS masih terbatas. Sebagian besar penelitian terdahulu masih terbatas pada variabel tunggal (*univariate*). Hal ini menunjukkan adanya celah penelitian yang dapat memberikan kontribusi baru, terutama dalam konteks pemodelan prediksi berbasis data lapangan dengan resolusi tinggi. Penelitian ini diharapkan mampu menjawab kebutuhan akan model prediksi yang tidak hanya akurat, tetapi juga efisien dan dapat diimplementasikan dalam sistem informasi maritim. Dengan demikian, hasil penelitian ini berpotensi memberikan manfaat langsung dalam mendukung keselamatan pelayaran, mitigasi risiko bencana, serta pengelolaan sumber daya kelautan.

2. METODE PENELITIAN

2.1. Alur Penelitian

Penelitian ini diawali dengan pengumpulan data, dilanjutkan dengan pra pemrosesan data, penerapan tiga algoritma prediksi (LSTM, *Random Forest*, dan XGBoost), hingga evaluasi performa model. Setiap tahap dilakukan dengan penyesuaian khusus agar sesuai dengan karakteristik data dan tujuan penelitian, yaitu memprediksi *multi-output* variabel atmosfer oseanografi secara akurat. Alur penelitian disajikan dalam bentuk diagram pada Gambar 1.



Gambar 1. Alur Penelitian

2.2. Pengumpulan Data

Pada penelitian ini, data yang digunakan berasal dari sistem AWS yang dimiliki oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). AWS berfungsi sebagai perangkat pemantau yang secara otomatis mencatat berbagai parameter meteorologi dan oseanografi, seperti kecepatan dan arah angin, suhu udara, ketinggian air, serta kelembaban relatif. Salah satu keunggulan utama dari data yang digunakan adalah resolusinya yang sangat tinggi, karena pencatatan dilakukan setiap satu jam tanpa jeda, mulai dari tahun 2022 hingga 2025. Setiap baris data merepresentasikan hasil pengamatan pada satu waktu tertentu, sehingga dalam sehari terkumpul 24 data untuk setiap variabel di satu lokasi. Data ini juga dilengkapi dengan informasi waktu pencatatan dan koordinat lokasi data. Sampel data yang digunakan dalam penelitian ini ditampilkan pada Tabel 1, yang menunjukkan struktur dataset serta contoh nilai dari masing-masing parameter yang dianalisis. Resolusi temporal yang tinggi ini menjadi aspek penting dalam mendukung analisis serta prediksi fenomena atmosfer dan oseanografi yang bersifat dinamis.

Tabel 1. Sampel Data AWS BMKG

no	waktu	lat	lon	winddir	windspeed	rh	watertemp	waterlevel
1	2022-01-01 00:00:00	-5.926459	105.886146	231.242 150	2.829114	86.5902 60	29.142172	3.152486
2	2022-01-01 00:00:00	-5.955795	105.937609	218.329 074	3.799139	87.9132 32	28.924046	2.335356
3	2022-01-01 00:00:00	-5.966479	105.914181	221.461 355	4.081405	89.7240 74	28.757494	2.399468
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
637 699 98	2025-02-06 23:58:00	-5.976946	105.951113	136.858 276	0.736269	82.1424 00	30.225185	1.499440
637 699 99	2025-02-06 23:58:00	-5.979169	105.940618	139.416 361	0.868192	81.9422 24	30.234263	1.461254
637 700 00	2025-02-06 23:58:00	-5.945715	105.905712	155.194 760	1.555569	81.4381 38	30.183579	1.648346

2.3. Pra-pemrosesan Data

Pra-pemrosesan merupakan tahap awal dalam menyiapkan data mentah sebelum dianalisis atau dimanfaatkan dalam pemodelan [9]. Pada tahap ini dilakukan ekstraksi fitur yang merupakan proses penting untuk mentransformasi data mentah menjadi fitur-fitur yang lebih signifikan [10]. Pada penelitian ini, variabel waktu dipecah menjadi tahun, bulan, hari, dan jam, serta ditambahkan representasi sinus dan cosinus pada atribut tersebut agar pola siklikal dan musiman dalam data dapat lebih mudah dikenali oleh model. Fitur tambahan seperti hari dalam minggu, indikator akhir pekan, dan musim juga disertakan karena terbukti berpengaruh terhadap dinamika atmosfer dan laut. Setelah itu, data dibersihkan dengan cara menghapus seluruh entri yang memiliki nilai kosong agar integritas analisis tetap terjaga [11].

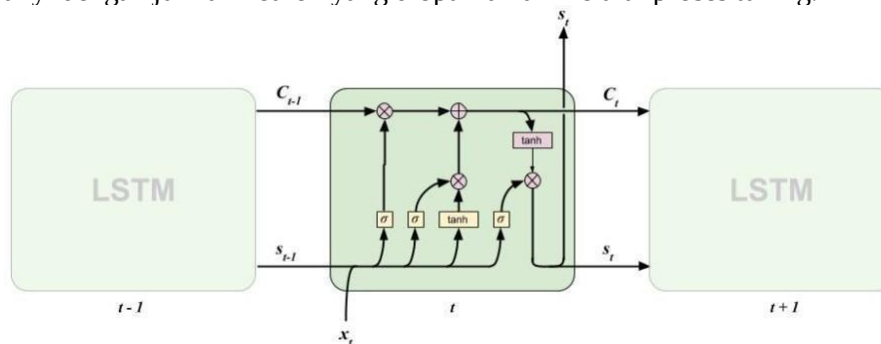
Selanjutnya, data dipisahkan menjadi bagian fitur (X) dan target (y), di mana fitur memuat parameter lingkungan dan waktu, sedangkan target terdiri dari kecepatan angin, tinggi muka air, suhu permukaan laut, kelembaban relatif, dan arah angin. Sebelum memasuki tahap pemodelan, dataset dipisahkan menjadi data pelatihan dan data pengujian untuk menghindari situasi di mana informasi dari data uji secara tidak sengaja telah digunakan selama proses pelatihan model [12]. Dalam penelitian ini data dibagi dengan proporsi 90% untuk pelatihan dan 10% untuk pengujian

agar evaluasi model berlangsung secara objektif.

Dalam tahap akhir pra pemrosesan, seluruh fitur numerik dinormalisasi menggunakan metode *StandardScaler* agar memiliki distribusi rata-rata nol dan standar deviasi satu, sehingga setiap fitur memiliki kontribusi yang seimbang dalam proses pelatihan [13]. Seluruh rangkaian pra pemrosesan ini diterapkan secara konsisten pada setiap algoritma yang digunakan dalam penelitian.

2.4. Long Short-Term Memory (LSTM)

Model LSTM dipilih karena mampu mengelola data deret waktu dengan ketergantungan jangka panjang melalui mekanisme sel internalnya, sehingga efektif dalam mempertahankan informasi historis pada data *time series* [14]. LSTM merupakan arsitektur RNN yang dirancang untuk mengatasi permasalahan *exploding* dan *vanishing gradient* saat memproses data dengan memori jangka panjang [15]. Arsitektur LSTM yang digunakan terdiri dari beberapa lapisan tersembunyi dengan jumlah neuron yang dioptimalkan melalui proses tuning.



Gambar 2. Arsitektur LSTM [16]

Berdasarkan Gambar 2, struktur LSTM mencakup tiga gerbang utama: forget gate yang berfungsi menghapus informasi yang tidak relevan dari memori sebelumnya, input gate yang mengatur masuknya informasi baru ke dalam memori, serta output gate yang menghasilkan keluaran berdasarkan kondisi memori terbaru. Ketiga gerbang tersebut berperan dalam menjaga stabilitas memori jangka panjang serta memungkinkan model untuk mempertahankan dan memperbarui informasi penting secara adaptif terhadap dinamika perubahan pada data deret waktu [17].

Model dilatih menggunakan data historis untuk memprediksi beberapa variabel target secara bersamaan. Data historis diolah menjadi urutan (*sequence*) sepanjang 12 langkah waktu, sehingga setiap input yang diberikan ke model LSTM berisi rangkaian data dari 12 waktu sebelumnya. Pendekatan ini memungkinkan LSTM untuk memahami dan menangkap pola hubungan antar waktu secara lebih mendalam. Arsitektur model terdiri dari beberapa lapisan tersembunyi, dengan jumlah neuron yang telah disesuaikan sehingga dapat menyeimbangkan antara kompleksitas model dan kemampuan generalisasi. Lapisan LSTM berfungsi untuk mengekstraksi fitur temporal dari data, sementara lapisan *dense* pada bagian akhir menghasilkan prediksi untuk seluruh variabel target secara bersamaan.

2.5. Random Forest Regression

Pemilihan algoritma *Random Forest Regression* pada penelitian ini didasarkan pada kemampuannya untuk memahami hubungan *non-linear* antara data *input* dengan target, serta keunggulannya dalam meminimalisir risiko *overfitting* saat menghadapi data yang kompleks [18]. *Random Forest Regression* adalah algoritma *machine learning* berbasis *supervised learning* yang memanfaatkan sekumpulan pohon keputusan untuk melakukan prediksi [19]. Setiap pohon regresi dalam model ini menghasilkan prediksi terhadap peubah respon, kemudian hasil prediksi dari seluruh pohon tersebut dirata-ratakan untuk memperoleh nilai akhir yang lebih akurat dan

stabil [20].

Model *Random Forest* dikonfigurasi melalui beberapa parameter utama, seperti jumlah pohon (*n_estimators*), jumlah minimal sampel untuk pemisahan node (*min_samples_split*), jumlah minimal sampel pada daun (*min_samples_leaf*), serta metode *bootstrap*. Penyesuaian parameter dilakukan dengan mempertimbangkan hasil validasi agar model yang diperoleh tidak mengalami *overfitting* dan mampu melakukan generalisasi dengan baik pada data baru. Implementasi *Random Forest* dilakukan dengan membungkusnya dalam *Multi Output Regression* agar dapat menangani prediksi *multi-output*, di mana seluruh variabel target diprediksi secara simultan.

2.6. *Extreme Gradient Boosting (XGBoost)*

XGBoost merupakan salah satu algoritma boosting yang populer karena keunggulannya dalam hal efisiensi komputasi dan akurasi prediksi pada berbagai permasalahan regresi [21]. XGBoost dikenal sebagai algoritma *tree boosting* yang sangat efisien dan banyak dimanfaatkan dalam berbagai aplikasi *machine learning* untuk menghasilkan prediksi nilai dari data [22]. Proses kerja XGBoost berlangsung secara iteratif, di mana setiap pohon keputusan yang dibentuk selanjutnya berfungsi untuk mengoreksi kesalahan prediksi yang dihasilkan oleh pohon-pohon sebelumnya, sehingga model menjadi semakin akurat seiring bertambahnya jumlah pohon [23][24].

Pada penelitian ini, XGBoost digunakan untuk melakukan prediksi multi-output, di mana model diharapkan mampu memprediksi beberapa variabel target secara bersamaan. Untuk mendukung kebutuhan tersebut, digunakan *wrapper multi output regressor* agar XGBoost dapat menangani lebih dari satu output dalam satu proses pelatihan. Model XGBoost kemudian dilatih menggunakan data latih yang telah dinormalisasi, dengan penyesuaian sejumlah parameter utama seperti jumlah pohon (*n_estimators*), kedalaman maksimum (*max_depth*), *learning rate*, serta parameter regularisasi (*reg lambda* dan *reg alpha*) agar menghasilkan model yang optimal dan tidak *overfitting*.

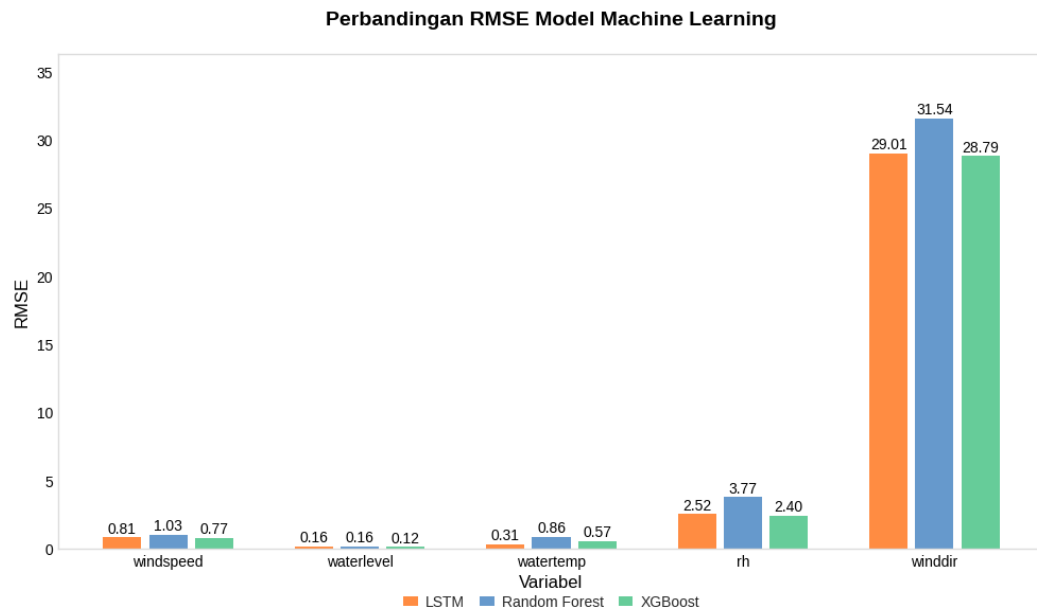
2.7. *Parameter Performa*

Evaluasi kinerja ketiga model dilakukan menggunakan beberapa metrik yang lazim digunakan pada analisis regresi, yaitu *Root Mean Squared Error (RMSE)* untuk mengukur tingkat kesalahan prediksi dan koefisien determinasi (R^2) untuk menilai tingkat kecocokan model terhadap data. Setiap metrik dihitung baik untuk masing-masing variabel target maupun secara rata-rata, sehingga efektivitas model dalam memprediksi multi-output dapat diukur secara menyeluruh. Hasil evaluasi ini menjadi dasar dalam menentukan model yang paling sesuai untuk prediksi variabel atmosfer oseanografi pada data AWS.

3. HASIL DAN PEMBAHASAN

3.1. *Analisis perbandingan nilai RMSE*

Evaluasi kinerja model berdasarkan nilai RMSE bertujuan untuk mengukur rata-rata kesalahan prediksi yang dihasilkan oleh model. Hasil perbandingan nilai RMSE antara model LSTM dan XGBoost ditampilkan dalam Gambar 3.



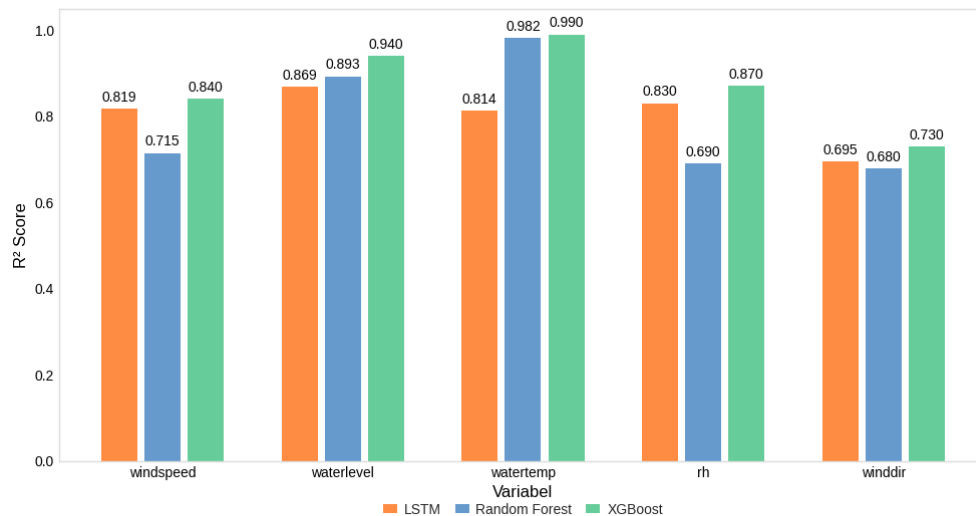
Gambar 3. Hasil Evaluasi RMSE pada Model LSTM, *Random Forest*, dan XGBoost

Hasil evaluasi menggunakan RMSE menunjukkan perbedaan performa yang cukup signifikan antara model LSTM, *Random Forest*, dan XGBoost. Secara umum, XGBoost menunjukkan nilai RMSE yang paling kecil pada empat dari lima parameter yang diuji, yaitu pada parameter *windspeed* sebesar 0,77, *waterlevel* 0,12, RH (*relative humidity*) 2,40, dan *winddir* 28,79. Nilai tersebut mengindikasikan bahwa XGBoost mampu meminimalkan kesalahan prediksi dan memberikan hasil yang lebih konsisten dibandingkan dengan dua model lainnya. Sementara itu, LSTM hanya menunjukkan keunggulannya pada satu parameter, yaitu *watertemp* sebesar 0,31. Sementara itu, model *Random Forest* menunjukkan nilai RMSE yang relatif lebih besar jika dibandingkan dengan XGBoost maupun LSTM.

Hasil evaluasi RMSE ini mengindikasikan bahwa XGBoost merupakan model yang efektif dalam menghasilkan prediksi dengan tingkat kesalahan minimal. Temuan ini sejalan dengan pernyataan [25] bahwa semakin kecil nilai RMSE, maka semakin rendah tingkat kesalahan prediksi yang dihasilkan, sehingga performa model menjadi lebih baik. Berdasarkan hal tersebut, dapat disimpulkan bahwa XGBoost lebih konsisten dalam memberikan prediksi dengan tingkat kesalahan minimal dibandingkan dengan LSTM dan *Random Forest*.

3.2. Analisis perbandingan nilai koefisien determinasi (R^2)

Evaluasi kinerja model berdasarkan nilai R^2 score bertujuan untuk mengukur sejauh mana model dapat menjelaskan variabilitas data pada setiap parameter yang diprediksi. R^2 score merupakan metrik evaluasi yang sangat penting dalam analisis regresi karena memberikan gambaran mengenai seberapa besar proporsi keragaman data yang mampu dijelaskan oleh model. Semakin besar nilai R^2 , maka semakin baik kemampuan model dalam menangkap pola dan hubungan yang terdapat dalam data. Hasil perbandingan nilai R^2 antara model LSTM, *Random Forest*, dan XGBoost ditampilkan dalam Gambar 4.



Gambar 4. Hasil Evaluasi R^2 pada Model LSTM, *Random Forest*, dan XGBoost

Berdasarkan hasil perbandingan nilai R^2 , terlihat bahwa performa ketiga model bervariasi pada setiap parameter. Untuk kecepatan angin (*windspeed*), XGBoost unggul dengan nilai 0.840, diikuti LSTM (0.819), dan *Random Forest* (0.715), menunjukkan bahwa XGBoost lebih mampu menangkap variabilitas pola angin. Pada tinggi muka air (*waterlevel*), ketiga model mencatat nilai tinggi di atas 0.86, dengan XGBoost (0.940) sedikit lebih unggul dibanding *Random Forest* (0.893) dan LSTM (0.869), mengindikasikan data ini relatif mudah diprediksi.

Berbeda pada suhu air (*watertemp*), XGBoost (0.990) dan *Random Forest* (0.982) menunjukkan akurasi hampir sempurna, jauh di atas LSTM (0.814), yang berarti model berbasis pohon lebih cocok untuk data ini. Untuk kelembaban relatif (RH), XGBoost tetap terdepan dengan 0.870, LSTM (0.830) menyusul, dan *Random Forest* (0.690) tertinggal cukup jauh, menegaskan keunggulan XGBoost dalam menangani pola atmosfer yang kompleks. Terakhir, pada arah angin (*winddir*), meskipun semua model memiliki R^2 lebih rendah (di bawah 0.75), XGBoost (0.730) masih menjadi yang terbaik, disusul LSTM (0.695) dan *Random Forest* (0.680). Secara keseluruhan, hasil evaluasi R^2 memperlihatkan bahwa XGBoost menunjukkan kemampuannya yang lebih baik dalam menjelaskan variabilitas data dibanding LSTM maupun *Random Forest*.

3.3. Analisis Kinerja Model Secara Komprehensif

Hasil evaluasi menggunakan RMSE dan R^2 menunjukkan perbedaan performa yang cukup jelas antara ketiga model. Secara umum, XGBoost menghasilkan nilai RMSE yang lebih rendah pada variabel *windspeed*, *waterlevel*, RH, dan *winddir*, serta nilai R^2 tertinggi pada semua variabel. Hal ini mengindikasikan bahwa XGBoost mampu menangkap hubungan non-linear dan interaksi kompleks antarvariabel dengan lebih konsisten, sehingga menghasilkan prediksi dengan kesalahan yang lebih kecil. Penelitian sebelumnya oleh [26] menunjukkan bahwa XGBoost efektif dalam memprediksi data cuaca yang kompleks, selaras dengan temuan pada penelitian ini.

Sementara itu, *Random Forest* menunjukkan nilai RMSE yang lebih tinggi pada semua variabel, menandakan performa yang kurang stabil. Nilai R^2 *Random Forest* lebih tinggi dibanding LSTM pada variabel ketinggian muka air dan suhu air, tetapi tetap kalah dibanding XGBoost pada semua variabel lainnya. Kondisi ini menegaskan keterbatasan *Random Forest* dalam memodelkan pola temporal dan interaksi non-linear yang kompleks. Sedangkan LSTM unggul khusus pada variabel suhu air, dengan RMSE paling rendah dibanding XGBoost maupun *Random Forest*. Selain itu, nilai R^2 LSTM juga lebih tinggi dibanding *Random Forest* pada variabel kecepatan angin, arah angin, dan kelembaban relatif, meskipun masih kalah dibanding XGBoost. Hal ini terjadi karena perubahan nilai suhu air bersifat gradual dan stabil seiring waktu

[27], sesuai dengan kemampuan memori jangka panjang LSTM yang mampu menangkap pola temporal gradual tersebut. Namun, pada variabel dengan dinamika tinggi dan fluktuatif, performa LSTM menurun, menunjukkan keterbatasannya dalam menangani variasi *non-linear* yang kompleks.

4. KESIMPULAN

Penelitian ini membandingkan kinerja tiga algoritma *machine learning*, yaitu LSTM, *Random Forest*, dan XGBoost dalam memprediksi *multi-output* variabel osean-atmosfer berbasis data AWS periode 2022–2025. Berdasarkan hasil evaluasi menggunakan RMSE dan R^2 , XGBoost menunjukkan performa terbaik dengan nilai RMSE terendah pada *windspeed* (0,77), *waterlevel* (0,12), RH (2,40), dan *winddir* (28,79), serta nilai R^2 tertinggi masing-masing sebesar 0,840; 0,940; 0,870; dan 0,730. Model LSTM memberikan hasil terbaik pada variabel *watertemp* dengan RMSE sebesar 0,31, meskipun nilai R^2 (0,814) masih berada di bawah XGBoost (0,990) untuk variabel yang sama. Sementara itu, *Random Forest* menunjukkan performa yang relatif lebih rendah dan kurang konsisten, dengan nilai R^2 terendah pada RH (0,690) dan *winddir* (0,680). Secara keseluruhan, XGBoost merupakan model yang paling efisien dan konsisten dalam menangani data multivariat dengan hubungan non-linear yang kompleks. Oleh karena itu, XGBoost direkomendasikan sebagai model utama dalam implementasi sistem prediksi variabel osean-atmosfer berbasis AWS guna mendukung keselamatan pelayaran dan mitigasi risiko bencana laut.

5. SARAN

Penelitian berikutnya juga dapat menambahkan variabel oseanografi lain, seperti arus laut atau salinitas, untuk memperkaya analisis hubungan multivariat antarparameter atmosfer dan laut. Selain itu, implementasi model ke dalam sistem berbasis web atau *dashboard* interaktif direkomendasikan agar hasil prediksi dapat dimanfaatkan secara *real-time* oleh pemangku kepentingan atau lembaga terkait dalam mendukung pengambilan keputusan operasional di bidang kemaritiman.

UCAPAN TERIMA KASIH

Penulis berterima kasih kepada BMKG Stasiun Meteorologi Maritim Kelas I Merak terutama dalam penyediaan data yang digunakan dalam penelitian ini. Ucapan terima kasih juga ditujukan kepada Bapak Willdan Aprizal Arifin, S.Pd., M.Kom., dan Bapak Dr. Herman Syafri, M.Pd., atas bimbingan, arahan, serta kontribusinya dalam penyelesaian dan penyempurnaan penelitian ini.

DAFTAR PUSTAKA

- [1] F. Pelu, “Dampak Perubahan Iklim Global Terhadap Parameter Oseanografi: Ancaman Bagi Ekosistem Laut dan Kehidupan Pesisir di Maluku,” *PERAUT: Jurnal Perikanan dan Kelautan*, vol. 2, no. 1, hlm. 38–46, 2025, doi: doi.org/10.70134/peraut.v2i1.592.
- [2] F. Setyanabi dan E. Mulyanie, “Analisis Dampak Perubahan Iklim Terhadap Geografi Maritim Indonesia,” *Jurnal Ilmiah Samudra Akuatika*, vol. 8, no. 1, hlm. 12–19, Jun 2024, doi: doi.org/10.33059/jisa.v8i1.9295.
- [3] N. N. B. Marscelina, I. G. L. Wijayakusuma, dan P. V. Swastika, “Perbandingan Metode LSTM dan TCN untuk Prediksi Gelombang Laut Berdasarkan Enam Parameter Oseanografi,” *JST (Jurnal Sains dan Teknologi)*, vol. 14, no. 1, hlm. 56–66, Mei 2025, doi: doi.org/10.23887/jstundiksha.v14i1.92590.

- [4] A. B. A. F. Adiaspoetri, “Analisis Bias Data Time Series dengan Algoritma Quantile Mapping (Studi Kasus : Presipitasi Kota Makassar),” Skripsi, Universitas Hasanuddin, Gowa, 2024.
- [5] A. Arwansyah, C. Susanto, dan N. Nurdiansah, “2Deep Model Prediksi Berbasis Weighting Average untuk Time Series Data 2Deep Weighting Average Based Prediction Model for Time Series Data,” *JURNAL SISFOTENIKA*, vol. 14, no. 2, hlm. 210–219, 2024, doi: doi.org/10.30700/sisfotenika.v14i2.462.
- [6] I. Botunac, J. Bosna, dan M. Matetić, “Optimization of Traditional Stock Market Strategies Using the LSTM Hybrid Approach,” *Information*, vol. 15, no. 3, Mar 2024, doi: doi.org/10.3390/info15030136.
- [7] H. A. Ramadhan, “Prediksi Curah Hujan Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average) (Studi Kasus: Kabupaten Sleman, Semarang, dan Surabaya),” Universitas Islam Indonesia, Yogyakarta, 2025.
- [8] H. A. Aziz dan H. A. Santoso, “Model Prediksi Stunting Anak di Indonesia Menggunakan Extreme Gradient Boosting,” *Jurnal Algoritma*, vol. 22, no. 1, hlm. 1072–1085, 2025, doi: doi.org/10.33364/algoritma/v.22-1.2289.
- [9] S. D. Prasetyo, S. S. Hilabi, dan F. Nurapriani, “Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naïve Bayes dan KNN,” *Jurnal KomtekInfo*, vol. 10, no. 1, hlm. 1–7, Jan 2023, doi: doi.org/10.35134/komtekinfo.v10i1.330.
- [10] H. Faiqoh, “Penerapan Ensemble Feature Selection untuk Mengurangi Dimensionalitas dalam Prediksi Data Time Series,” Skripsi, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Tangerang Selatan, 2023.
- [11] A. Sispianygala, S. S. Berutu, dan J. Jatmitka, “Pengembangan Aplikasi Sistem Rekomendasi Tempat Wisata dengan Collaborative Filtering,” *Progresif: Jurnal Ilmiah Komputer*, vol. 20, no. 2, hlm. 828–838, 2024, doi: dx.doi.org/10.35889/progresif.v20i2.2044.
- [12] Y. Setiawan, “Data Mining berbasis Nearest Neighbor dan Seleksi Fitur untuk Deteksi Kanker Payudara,” *Jurnal Informatika: Jurnal pengembangan IT (JPIT)*, vol. 8, no. 2, hlm. 89–96, 2023, doi: doi.org/10.30591/jpit.v8i2.4994.
- [13] A. Syaputra, “Implementation of Machine Learning Algorithms for Predicting Student Final GPA Using Multiclass Classification Models,” *Journal of Artificial Intelligence and Software Engineering*, vol. 5, no. 2, hlm. 660–675, 2025, doi: dx.doi.org/10.30811/jaise.v5i2.7015.
- [14] T. Lattifia, P. W. Buana, dan N. K. D. Rusjyanthi, “Model Prediksi Cuaca Menggunakan Metode LSTM,” *JITTER-Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 1, 2022.
- [15] A. Tholib, N. K. Agusmawati, dan F. Khoiriyah, “Prediksi Harga Emas Menggunakan Metode LSTM dan GRU,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 11, no. 3, hlm. 620–627, Agu 2023, doi: doi.org/10.23960/jitet.v11i3.3250.
- [16] I. K. A. Enriko, F. N. Gustiyana, dan R. H. Putra, “Komparasi Hasil Optimasi Pada Prediksi Harga Saham PT. Telkom Indonesia Menggunakan Algoritma Long Short Term Memory,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 2, hlm. 659–667, Apr 2023, doi: doi.org/10.30865/mib.v7i2.5822.
- [17] A. B. Fawait, M. Jamil, S. Rahmah, dan S. Sugiarto, “Penerapan Metode LSTM untuk Prediksi Harga Ethereum,” *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 9, no. 3, hlm. 243–252, 2025, doi: dx.doi.org/10.30872/jurti.v9i3.22169.
- [18] D. N. Handayani dan S. Qutub, “Penerapan Random Forest untuk Prediksi dan Analisis Kemiskinan,” *RIGGS: Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, hlm. 405–412, Mei 2025, doi: doi.org/10.31004/riggs.v4i2.512.

- [19] E. Fitri, “Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah,” *Journal of Applied Computer Science and Technology (JACOST)*, vol. 4, no. 1, hlm. 58–64, 2023, doi: doi.org/10.52158/jacost.491.
- [20] E. Fitri dan D. Riana, “Analisa Perbandingan Model Prediction dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression dan Multilayer Perceptron,” *METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi*, vol. 6, no. 1, hlm. 69–78, 2022, doi: doi.org/10.46880/jmika.Vol6No1.pp69-78.
- [21] B. W. Sari dan D. Prabowo, “Analisis Perbandingan Prediksi Harga Rumah dengan Random Forest, Gradient Boosting, dan XGBoost,” *Intellect : Indonesian Journal of Innovation Learning and Technology*, vol. 04, no. 01, hlm. 42–51, 2025, doi: doi.org/10.57255/intellect.v4i1.1385.
- [22] M. A. Rayadin, M. Musaruddin, R. A. Saputra, dan I. Isnawaty, “Implementasi Ensemble Learning Metode XGBoost dan Random Forest untuk Prediksi Waktu Penggantian Baterai Aki,” *BIOS : Jurnal Teknologi Informasi dan Rekayasa Komputer*, vol. 5, no. 2, hlm. 111–119, Agu 2024, doi: doi.org/10.37148/bios.v5i2.128.
- [23] A. Joshi, C. Vishnu, C. K. Mohan, dan B. Raman, “Application of XGBoost model for early prediction of earthquake magnitude from waveform data,” *Journal of Earth System Science*, vol. 133, no. 5, Mar 2024, doi: doi.org/10.1007/s12040-023-02210-1.
- [24] Z. Gustiyandi, A. Aradea, dan H. Sulastri, “Prediksi Penjualan Produk Menggunakan Algoritma Xtreme Gradient Boosting,” *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 9, no. 1, hlm. 20–29, 2025, doi: dx.doi.org/10.30872/jurti.v9i1.20419.
- [25] I. Amansyah, J. Indra, E. Nurlaelasari, dan A. R. Juwita, “Prediksi Penjualan Kendaraan Menggunakan Regresi Linear Studi Kasus pada Industri Otomotif di Indonesia,” *INNOVATIVE: Journal of Social Science Research*, vol. 4, no. 4, hlm. 1199–1216, 2024, doi: doi.org/10.31004/innovative.v4i4.12735.
- [26] A. Syahreza, N. K. Ningrum, dan M. A. Syahrazy, “Perbandingan Kinerja Model Prediksi Cuaca: Random Forest, Support Vector Regression, dan XGBoost,” *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 2, hlm. 526–534, Des 2024, doi: doi.org/10.29408/edumatic.v8i2.27640.
- [27] A. Alwi, “Peramalan Suhu Udara di Kabupaten Malang Menggunakan Long Short-Term Memory (LSTM),” Skripsi, Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang, 2025.