Jurnal Informatika
Vol. x, No. x, bulan 20xx, pp.1-10
P-ISSN:1693-3877 / E-ISSN: 2407-1544
DOI:10.30873

Evaluasi Kinerja Model Deep Learning dalam Memprediksi Kejadian Hujan Di Wilayah Panjang Bandar Lampung

Tarjono[#], Joko Triloka[*], Suci Mutiara[#]

[#] *Fakultas Ilmu Komputer, Institut Informatika Dan Bisnis Darmajaya, Jl, Z.A. Pagar Alam No.93, Bandar Lampung, 35141, Indonesia*
***** *Fakultas Ilmu Komputer, Institut Informatika Dan Bisnis Darmajaya, Jl, Z.A. Pagar Alam No.93, Bandar Lampung, 35141, Indonesia*
E-mail: : tarjono.2321210020p@mail.darmajaya.ac.id, joko.triloka@darmajaya.ac.id, sucimutiara@darmajaya.ac.id

Abstract — **Global warming and climate change have increased the frequency and intensity of extreme weather events, significantly impacting human life and the environment. Urban areas such as Kecamatan Panjang in Bandar Lampung City frequently experience flooding due to extreme rainfall and poor drainage systems. This study compares the effectiveness of three deep learning model architectures- Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Transformers — in predicting rainfall events in Kecamatan Panjang. The data used includes key meteorological variables such as air temperature, dew point, humidity, and air pressure, collected from the Maritime Meteorology Station in Panjang (BMKG) over the past three years. The models were trained using this historical data, with the data divided into training and testing sets. The study results indicate that the Transformer model performs best with the highest accuracy compared to CNN and RNN. The Transformer model efficiently captures long-term dependencies in sequential data, providing more accurate and timely predictions. Model performance evaluation was conducted using accuracy, F1 score, precision, recall, ROC AUC, RMSE, and MAE metrics. The use of deep learning models in rainfall prediction is expected to assist in flood risk mitigation and planning for adaptation to increasingly frequent extreme weather due to climate change. This research significantly advances more accurate and efficient weather prediction systems for urban areas prone to hydrological disasters.**

Key word **— deep learning; flood mitigation; global warming; rainfall prediction.**

Abstrak — **Pemanasan global dan perubahan iklim telah meningkatkan frekuensi dan intensitas cuaca ekstrem yang berdampak signifikan pada kehidupan manusia dan lingkungan. Wilayah perkotaan seperti Kecamatan Panjang di Kota Bandar Lampung sering mengalami banjir akibat curah hujan ekstrem dan buruknya sistem drainase. Penelitian ini bertujuan untuk membandingkan efektivitas tiga arsitektur model deep learning, yaitu Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), dan Transformer, dalam memprediksi kejadian hujan di Kecamatan Panjang. Data yang digunakan mencakup variable meteorologis utama, seperti suhu udara, titik embun, kelembapan udara, dan tekanan udara, yang dikumpulkan dari Stasiun Meteorologi Maritim Panjang (BMKG) selama tiga tahun terakhir. Model dilatih menggunakan data historis ini dengan pembagian data menjadi set pelatihan dan pengujian. Hasil penelitian menunjukkan bahwa model Transformer memiliki kinerja terbaik dengan akurasi tertinggi dibandingkan CNN dan RNN. Model Transformer mampu menangkap hubungan jangka panjang dalam data sekuensial dengan efisien, memberikan prediksi yang lebih akurat dan tepat waktu. Evaluasi kinerja model dilakukan menggunakan metrik seperti akurasi, F1 score, precision, recall, ROC AUC, RMSE, dan MAE. Penggunaan model deep learning dalam prediksi hujan diharapkan dapat membantu dalam mitigasi risiko banjir dan perencanaan adaptasi terhadap cuaca ekstrem yang semakin sering terjadi akibat perubahan iklim. Penelitian ini memberikan kontribusi signifikan dalam pengembangan sistem prediksi cuaca yang lebih akurat dan efisien untuk wilayah perkotaan yang rentan terhadap bencana hidrologi.**

Kata kunci— **deep learning; mitigasi banjir; pemanasan global; prediksi hujan.**

I. PENDAHULUAN

Kecamatan Panjang, yang merupakan bagian dari Ibu Kota Provinsi Lampung, memiliki histori curah hujan yang tinggi dan sering mengalami perubahan cuaca yang tidak terduga. Ketidaktepatan dalam memprediksi kejadian hujan di wilayah ini dapat berdampak serius terhadap berbagai sektor, termasuk mobilitas masyarakat, aktivitas penyeberangan, pertanian, serta pengelolaan sumber daya air. Tanpa prediksi hujan yang akurat, risiko bencana seperti banjir dan longsor meningkat, yang dapat

menyebabkan kerusakan infrastruktur, kehilangan hasil pertanian, serta mengganggu aktivitas ekonomi dan sosial masyarakat. Misalnya, banjir yang terjadi di beberapa bagian Kota Bandar Lampung pada tahun-tahun sebelumnya dan beberapa bulan yang lalu mengakibatkan kemacetan parah, kerusakan rumah warga, serta terhambatnya distribusi logistik. Selain itu, longsor yang dipicu oleh curah hujan tinggi di wilayah perbukitan sekitar Lampung telah menyebabkan gangguan akses jalan dan bahkan korban jiwa. Sebagai contoh di Kecamatan Panjang di Bandar Lampung mengalami beberapa kejadian banjir yang signifikan. Pada pertengahan Januari 2025, hujan deras menyebabkan banjir besar yang melanda 17 wilayah di Kota Bandar Lampung, termasuk Kecamatan Panjang. Banjir ini mengakibatkan lebih dari 14.000 rumah terdampak, dengan 2.880 rumah di Kecamatan Panjang terkena imbasnya (https://www.kompas.id/artikel/banjir-di-bandar-lampung-lebih-dari-14000-rumah-terdampak). Sebelumnya, pada akhir September 2024, hujan lebat mengakibatkan genangan air setinggi lutut di beberapa area, seperti Kampung Rawa Laut di Kecamatan Panjang. Banjir ini menyebabkan kemacetan lalu lintas dan mengganggu aktivitas masyarakat setempat (https://lampungpro.co/news/hujan-deras-rajabasa-sukarame-panjang-dan-sukabumi-bandar-lampung-kembali-dikepung-banjir).

Kejadian-kejadian ini menekankan pentingnya prediksi hujan yang akurat untuk mengurangi risiko banjir di masa mendatang. Pengembangan model prediksi hujan yang lebih akurat menjadi sangat penting untuk mendukung upaya mitigasi bencana dan meningkatkan kesiapsiagaan dalam menghadapi dampak cuaca ekstrem di wilayah ini.

Dalam menghadapi tantangan ini, pemanfaatan teknologi berbasis kecerdasan buatan (Artificial Intelligence/AI) menjadi solusi inovatif yang dapat meningkatkan akurasi prediksi hujan. Model pembelajaran mesin seperti Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), dan Transformer telah terbukti mampu menganalisis pola curah hujan berdasarkan data historis dan variabel meteorologi lainnya. Dengan AI, sistem prediksi cuaca dapat mengolah data dalam jumlah besar secara real-time, memberikan hasil yang lebih cepat dan akurat dibandingkan metode konvensional. Implementasi teknologi ini dapat membantu pemerintah daerah, instansi terkait, serta masyarakat dalam mengambil tindakan preventif yang lebih efektif yaitu peringatan dini banjir sebelum bencana terjadi. Dengan demikian, integrasi kecerdasan buatan dalam sistem prediksi cuaca dapat menjadi langkah strategis dalam mengurangi risiko bencana akibat hujan ekstrem di Kecamatan Panjang dan wilayah sekitarnya.

Penelitian sebelumnya telah menggunakan Convolutional Neural Netwroks (CNN) dalam mengidentifikasi pola-pola hujan dengan lebih baik melalui analisis citra satelit dan peta cuaca, dan juga mampu memberikan prediksi yang lebih akurat dibandingkan dengan metode konvensional [1]. Selain itu, menurut [2] CNN mampu mengolah data meteorologi dengan efisien dan menghasilkan model prediksi yang tangguh terhadap variasi cuaca terutama dalam memprediksi kejadian hujan.

Penelitian lainnya menggunakan Recurrent Neural Network (RNN) untuk memprediksi kejadian hujan. Penelitian yang dilakukan oleh [3] menjelaskan bahwa model RNN dapat digunakan pada data urutan historis yang panjang untuk memprediksi dinamika temporal cuaca yang memiliki korelasi terhadap prediksi kejadian hujan. Sedangkan menurut [4] menunjukkan bahwa LSTM sebagai model prediksi lanjutan dari RNN mampu untuk mempertahankan informasi yang relevan dalam jangka yang lebih panjang untuk meningkatkan performa prediksi cuaca.

Sedangkan [5] mempopulerkan pertama kali model Transformer digunakan dalam bidang meteorologi untuk memprediksi cuaca. Menurut penelitian yang dilakukan oleh [6] menunjukkan bahwa hasil prediksi menggunakan Transformer memberikan hasil yang menjanjikan dalam prediksi cuaca dengan kemampuannya menangani hubungan jangka panjang antar data cuaca yang dilatih. Transformer bekerja menggunakan mekanisme perhatian (attention mechanism) untuk memfokuskan mencari informasi yang penting dalam suatu kumpulan data.

Penelitian ini sendiri bertujuan untuk membandingkan performa CNN, RNN, dan Transformer dalam memprediksi kejadian hujan di wilayah Kecamatan Panjang, Bandar Lampung. Dengan menggunakan dataset meteorologi yang komprehensif, penelitian ini akan mengevaluasi keakuratan

mode, efisiensi, dan kemampuan generalisasi dari ketiga model yang digunakan. Sebagaimana disampaikan oleh [7] bahwa pemilihan model yang tepat dapat secara signifikan meningkatkan kualitas dari hasil prediksi yang dilakukan, dan menurut [8] menyarankan agar integrasi menggunakan berbagai model deep learning untuk memperoleh hasil penelitian yang lebih optimal. Hasil dari penelitian ini akan memperkaya wawasan bagi pengembangan sistem prediksi cuaca agar menjadi lebih handal dan akurat.

II. METODE

A. Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini yaitu data cuaca yang dikumpulkan setiap jam oleh pengamat cuaca di BMKG Maritim Lampung mencakup parameter-parameter penting seperti suhu udara, titik embun, kelembapan udara, dan tekanan udara. Proses pengumpulan ini dilakukan secara manual dan sistematis untuk memastikan data yang diperoleh akurat dan konsisten menggunakan alat pengamatan yang modern. Data ini kemudian disusun dalam bentuk tabel dan database, yang akan digunakan sebagai input untuk model prediksi cuaca. Jumlah data yang terkumpul untuk pelatihan dan pengujian model berjumlah 7.300 data, yang merupakan kumpulan data cuaca dari berbagai waktu dan kondisi. Data target sebanyak 730 digunakan untuk mengukur hasil prediksi model dengan kejadian nyata, sedangkan 1.650 data validasi digunakan untuk memvalidasi model di luar data pelatihan dalam menghindari overfitting. Pengamatan dilakukan dengan cermat dan mematuhi standar prosedur operasional yang ditetapkan oleh BMKG.

B. Waktu dan Lokasi Penelitian

Penelitian ini menggunakan data cuaca dari periode 2022-2023 sebagai data pelatihan dan pengujian. Pemilihan periode ini didasarkan pada ketersediaan data yang cukup serta keragaman kondisi cuaca yang terjadi dalam dua tahun tersebut. Keragaman ini akan memaksimalkan model untuk belajar dari berbagai kondisi cuaca, termasuk kondisi ekstrem seperti badai atau panas terik, yang penting untuk meningkatkan kemampuan prediksi model.Data dari tahun 2024 digunakan sebagai data validasi. Sedangkan untuk Lokasi penelitian dilakukan di stasiun BMKG Maritim Panjang, Pelabuhan Panjang, Kota Bandar Lampung.

C. Data Penelitian

Data penelitian mencakup berbagai parameter cuaca yang diukur pada dua waktu utama dalam sehari, yaitu jam 13.00 WIB atau jam 06.00 UTC, dan jam 16.00 WIB atau jam 09.00 UTC. Pemilihan waktu ini didasarkan pada variasi suhu harian yang biasanya terjadi, di mana suhu dan kondisi atmosferik lainnya dapat berubah secara signifikan antara pagi, siang, dan sore hari. Data ini penting untuk menangkap dinamika cuaca harian dan membantu dalam membuat prediksi yang lebih tepat. Suhu udara diukur pada dua waktu tersebut untuk melihat perbedaan temperatur yang terjadi, yang bisa menjadi indikasi perubahan cuaca. Titik embun dan kelembapan udara juga diukur karena parameter ini berhubungan erat dengan kemungkinan terjadinya hujan atau kondisi cuaca lainnya. Tekanan udara, baik QFF maupun QFE, diukur untuk memberikan indikasi adanya sistem tekanan rendah atau tinggi yang dapat mempengaruhi cuaca lokal. Selain parameter fisik tersebut, data kondisi cuaca seperti hujan atau tidaknya (dengan nilai 0 untuk tidak hujan dan 1 untuk hujan) juga dikumpulkan. Informasi ini sangat penting karena dapat digunakan sebagai variabel target dalam model prediksi cuaca. Dengan memiliki data yang lengkap dan variatif, model prediksi cuaca yang dikembangkan dapat diharapkan memberikan hasil yang akurat dan dapat diandalkan untuk mendukung berbagai aktivitas di pelabuhan. Contoh data sebagai inputan untuk model prediksi ditunjukkan pada Tabel 1.

Tabel 1. Sampel Data Inputan Model Prediksi

Tanggal	t06	t09	td06	td09	rh06	rh09	qff06	qff09	qfe06	qfe09	hujan
1-Jan-22	30.8	31.4	25.6	25.8	74.1	72.5	1010.1	1010.1	1009.8	1007	0
2-Jan-22	33.6	31.8	25.8	24.4	64.1	65.3	1009.5	1007.6	1009.2	1007.3	0
3-Jan-22	31	30.8	26.2	25.4	76.2	73.5	1009.8	1008	1009.5	1007.7	0
4-Jan-22	32	32.4	25.8	26.8	70.3	73	1010.6	1008.5	1010.3	1008.2	0
5-Jan-22	32.8	32.6	25.2	27	64.8	73.1	1009.3	1007.9	1009	1007.6	1

Data yang telah diproses kemudian digunakan untuk pelatihan model dengan tiga metode: Convolutional Neural Networks (CNN), Recurrent Neural Network (RNN), dan Transformer.

Evaluasi kinerja model dilakukan menggunakan berbagai metrik, termasuk Akurasi (Test), F1-Score, Precision, Recall, ROC AUC, RMSE, MAE, Korelasi (Train), Korelasi (Test), Selisih Korelasi, Rata-Rata Error (Train), dan Akurasi (Train). Model dengan nilai akurasi < 70 akan diklasifikasikan menjadi Kurang Baik dan untuk model dengan nilai akurasi > 70% akan diklasifikasikan menjadi Baik. Jika semua memiliki status Kurang Baik maka akan dilihat model terbaik dari semua yang Kurang Baik berdasarkan hasil evaluasinya.

Berikut formula yang digunakan untuk masing-masing metrik evaluasi yang digunakan dalam mengukur kinerja dan performa model:

1. Akurasi (Test/Train)

Metrik ini mengukur persentase prediksi yang benar dibandingkan dengan jumlah keseluruhan sampel (total) dengan persamaan yaitu:

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Jumlah\ Keseluruhan\ Sampel}\ x\ 100 \quad (1)$$

2. Precision

Metrik ini untuk mengukur seberapa akurat prediksi positif dari model dengan cara membandingkan jumlah prediksi positif yang benar terhadap jumlah keseluruhan prediksi positif menggunakan formula:

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (2)$$

3. Recall

Metrik ini mengukur seberapa baik model yang digunakan dapat mendeteksi semua data positif dengan cara membandingkan jumlah prediksi positif yang benar terhadap total data positif yang sebenarnya menggunakan formula:

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad (3)$$

4. F1-score

Metrik evaluasi untuk menghitung nilai rata-rata harmonis antara nilai Precision dan Recall agar terjadi keseimbangan diantara keduanya menggunakan formaula:

$$F1 = 2\ x\ (Precision\ x\ Recall)/(Precision + Recall) \quad (4)$$

5. Area Under the Curve (AUC).

Digunakan untuk mengukur peforma model klasifikasi dengan cara menganalisis trade-off antara True Positives Rate (TPR) dan False Positives Rate (FPR) pada berbagai treshold atau ambang batas menggunakan formula:

$$AUC = \sum_{i=1}^{n-1}(FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1}+TPR_1}{2} \quad (5)$$

6. Root Mean Squared Error (RMSE)

Digunakan untuk mengukur sejauh mana model prediksi yang digunakan menyimpang dari nilai yang sebenarnya dengan cara menghitung akar kuadrat dari rata-rata selisih kuadrat antara nilai aktual dan juga nilai prediksi menggunakan formula:

$$RMSE = \sqrt{\frac{\sum_{i-1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (6)$$

7. Mean Absolute Error (MAE)

Metrik ini mengukur rata-rata selisih absolut antara nilai aktual dan prediksi tanpa perlu memperhitungkan arah kesalahan menggunakan formula:

$$MAE = \frac{\sum_{i=1}^{n}|\hat{y}_i - y_i|}{n} \quad (7)$$

8. Korelasi (Train/Test)

Metrik untuk mengukur sejauh mana 2 variabel saling berhubungan secara linear yaitu nilai aktual dan juga nilai prediksi menggunakan formula:

$$Korelasi = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sqrt{\sum_{-=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}} \quad (8)$$

9. Selisih Korelasi

Metrik selisih korelasi digunakan untuk mengukur perbedaan yang muncul antara kekuatan hubungan liniear antara nilai aktual dan prediksi pada saat pelatihan dan pengujian menggunakan formula:

$$Selisih\ Korelasi = |Korelasi\ (Train) - Korelasi\ (Test)| \quad (9)$$

10. Rata-Rata Error (Train)

Metrik rata-rata error digunakan untuk mengukur rata-rata selisih antara nilai prediksi dan nilai aktual pada data pelatihan menggunakan formula:

$$Rata - Rata\ Error = \frac{\sum_{i=1}^{n}(\bar{y} - y_i)}{n} \quad (10)$$

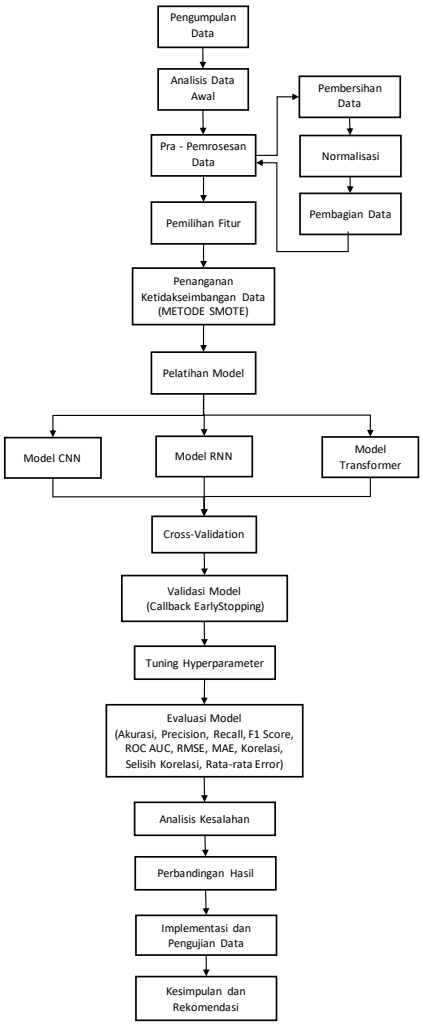
D. Diagram Alir Penelitian

Gambar 1. Menunjukkan tahapan peneltian mulai dari input data hingga menghasilkan prediksi. Secara garis besar, tahapan penelitian dimulai dari pengumpulan data yang meliputi data-data suhu udara, titik embun, kelembapan udara, QFF, dan QFE. Data-data ini bersumber dari hasil pengamatan yang dilakukan oleh petugas observasi di BMKG Maritim Panjang Bandar Lampung. Sebelum diolah dengan model prediksi data unsur cuaca ini akan melalui proses preprocessing, yaitu data-data akan melewati proses normalisasi untuk memastikan nilai-nilainya berada pada rentang tertentu atau nilai

yang telah ditetapkan. Kemudian dilakukan proses reshaping untuk menyesuaikan format data dengan kebutuhan model yang digunakan, serta proses ekstraksi fitur agar model dapat fokus pada elemen penting dari data cuaca yang diinginkan. Tahap berikutnya dilakukan penentuan fitur yang paling relevan untuk meningkatkan performa model.

Proses selanjutnya dilakukan penanganan ketidakseimbangan data dengan teknik Synthetic Minority Over-sampling Technique (SMOTE) untuk menyeimbangkan distribusi kelas dalam dataset. Tahap berikutnya yaitu melatih model menggunakan tiga arsitektur deep learning: CNN, RNN, dan Transformer, setelah pelatihan model selanjutnya dilakukan validasi silang (Cross Validation) untuk memastikan model tidak mengalami overfitting dan memiliki performa yang konsisten. Teknik Callback Early Stopping kemudian diterapkan untuk menghentikan pelatihan model secara otomatis saat tidak ada peningkatan performa lebih lanjut untuk selanjutnya dilakukan penyesuaian parameter model (Tuning Hyperparameter) untuk mengoptimalkan kinerjanya.

Tahap berikutnya melakukan evaluasi model menggunakan berbagai metrik yaitu akurasi, recall, presisi, F1-score, AUC, RMSE, MAE, korelasi, selisih korelasi, dan rata-rata error. Identifikasi kelemahan model kemudian dilakukan untuk meninjau prediksi yang salah. Selanjutnya adalah perbandingan hasil performa ketiga model untuk menentukan metode terbaik. Dari pemilihan metode terbaik, selanjutnya dilakukan implementasi dan pengujian data pada data baru untuk menguji kinerjanya di dunia nyata. Hasil temuan penelitian kemudian dirangkum untuk pengembangan lebih lanjut.



Gambar 1. Alur Penelitian

P-ISSN:1693-3877 / E-ISSN: 2407-1544
DOI:10.30873

III. HASIL DAN PEMBAHASAN

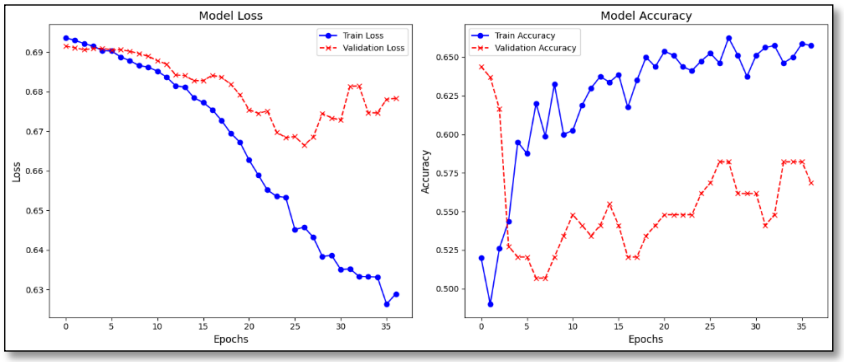
A. Model Prediksi Hujan Convolutional Neural Networks (CNN)

Gambar 2. menunjukkan dua grafik yang menggambarkan model Loss dan model Accuracy selama proses pelatihan model CNN. Grafik kiri menampilkan perubahan loss untuk data pelatihan (train Loss) dan validasi (validation loss) seiring bertambahnya jumlah epochs, sedangkan grafik kanan menunjukkan akurasi model (train accuracy dan validation accuracy) pada data pelatihan dan validasi selama pelatihan.

Pada grafik model Loss (kiri), terlihat bahwa loss pada data pelatihan (garis biru) secara konsisten menurun seiring bertambahnya epochs, yang menunjukkan bahwa model semakin mampu meminimalkan kesalahan prediksi pada data pelatihan. Sebaliknya, loss pada data validasi (garis merah) menunjukkan fluktuasi setelah beberapa epoch awal, dengan trend yang tidak menurun secara signifikan setelah epoch ke-15. Hal ini mengindikasikan meskipun model semakin baik dalam menangani data pelatihan, kemampuan generalisasinya terhadap data baru (validasi) tidak meningkat dengan signifikan, yang dapat menjadi indikasi overfitting.

Pada grafik model Accuracy (kanan), akurasi pada data pelatihan (garis biru) meningkat secara bertahap, terutama setelah sekitar 10 epoch, mencapai lebih dari 60%. Namun, akurasi pada data validasi (garis merah) menunjukkan fluktuasi yang signifikan, dan bahkan cenderung stagnan atau menurun pada beberapa titik setelah awal pelatihan. Fenomena ini menunjukkan bahwa meskipun model mampu belajar dari data pelatihan, model tidak mampu mempertahankan atau meningkatkan performanya pada data yang belum pernah dilihat (validasi), yang juga merupakan indikasi overfitting.

Kedua grafik tersebut menunjukkan adanya overfitting pada model CNN yang dilatih. Hal ini dapat dilihat dari perbedaan yang semakin besar antara loss dan akurasi pada data pelatihan dan validasi setelah beberapa epoch. Meskipun akurasi pada data pelatihan meningkat dan loss menurun, performa model pada data validasi tidak menunjukkan peningkatan yang sama.



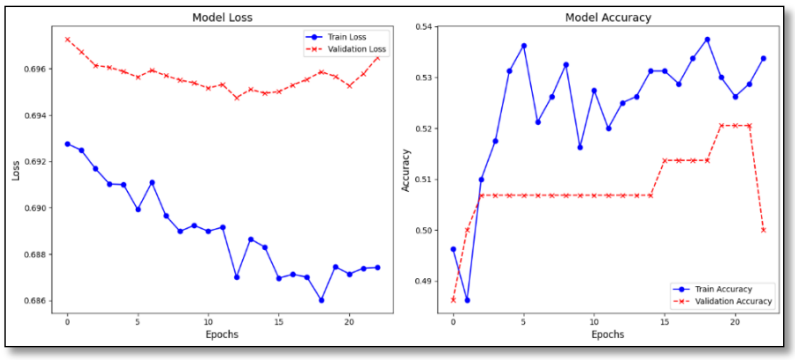
Gambar 2. Model *Loss* dan Model *Accuracy* CNN

B. Model Prediksi Hujan Recurrent Neural Network (RNN)

Gambar 3. menunjukkan dua grafik yang menampilkan performa dari model Recurrent Neural Network (RNN) selama proses pelatihan, yaitu model *Loss* dan model *Accuracy*. Grafik di sebelah kiri menggambarkan model *loss* yang menunjukkan perubahan nilai *loss* atau kesalahan pada model selama proses pelatihan, sementara grafik di sebelah kanan menggambarkan model *Accuracy* yang

P-ISSN:1693-3877 / E-ISSN: 2407-1544
DOI:10.30873

menunjukkan akurasi model. Kedua grafik ini membantu dalam memahami bagaimana model belajar dari data yang diberikan.



Gambar 3. Model *Loss* dan Model *Accuracy* RNN

Pada grafik model Loss, terlihat bahwa nilai loss pada data pelatihan (train loss) terus mengalami penurunan seiring bertambahnya jumlah epoch, yang menandakan bahwa model semakin baik dalam meminimalkan kesalahan prediksi pada data pelatihan. Namun, yang menjadi perhatian adalah nilai loss pada data validasi (validation loss) yang cenderung stagnan atau bahkan sedikit meningkat setelah beberapa epoch. Sama seperti CNN, hal ini mengindikasikan meskipun model terus belajar dan menjadi lebih baik pada data pelatihan, performanya pada data yang tidak terlihat tidak meningkat secara signifikan dan dapat menunjukkan tanda-tanda overfitting setelah beberapa epoch tertentu.

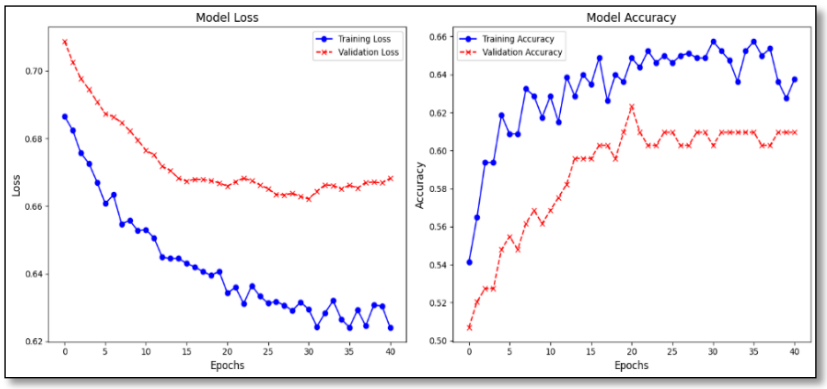
Grafik model Accuracy di sebelah kanan menunjukkan pola yang sedikit berbeda. Pada awal pelatihan, akurasi pada data pelatihan (train accuracy) meningkat secara signifikan, namun setelah beberapa epoch, peningkatannya menjadi lebih lambat dan berfluktuasi. Akurasi pada data validasi (validation accuracy), di sisi lain, cenderung stabil pada nilai yang relatif rendah setelah beberapa epoch pertama dan tidak menunjukkan peningkatan yang berarti. Bahkan, pada epoch terakhir, terjadi penurunan yang drastis pada akurasi validasi, yang bisa menjadi tanda overfitting atau bahwa model tidak generalisasi dengan baik pada data baru.

Perbedaan antara akurasi dan loss pada data pelatihan dan data validasi ini menunjukkan bahwa model mungkin sudah mulai overfitting, yaitu ketika model terlalu menyesuaikan diri dengan data pelatihan sehingga kehilangan kemampuannya untuk generalisasi pada data baru. Overfitting ini sering terjadi ketika model terlalu kompleks atau ketika pelatihan dilakukan terlalu lama tanpa mekanisme penghentian dini (early stopping). Dalam kasus ini, meskipun model semakin akurat dalam memprediksi data pelatihan, performanya pada data validasi justru memburuk atau tidak meningkat.

C. Model Prediksi Hujan Transformer

Model prediksi yang ketiga diimplementasikan menggunakan arsitektur Transformer yang bertujuan untuk memprediksi curah hujan panjang berdasarkan data cuaca historis. Gambar 4. menampilkan dua grafik yang menggambarkan kinerja model Transformer dalam tugas prediksi melalui dua metrik utama: model Loss dan model Accuracy. Grafik pertama di sebelah kiri memperlihatkan bagaimana nilai loss, baik untuk data pelatihan (training loss) maupun data validasi (validation loss), berubah seiring dengan bertambahnya jumlah epoch selama proses pelatihan. Grafik kedua di sebelah kanan menunjukkan perubahan akurasi model, baik untuk data pelatihan (training accuracy) maupun data validasi (validation accuracy), sepanjang epoch yang sama.

P-ISSN:1693-3877 / E-ISSN: 2407-1544
DOI:10.30873



Gambar 4. Model *Loss* dan Model *Accuracy* Transformer

Pada grafik model Loss, dapat dilihat bahwa loss pada data pelatihan terus menurun secara konsisten, menunjukkan bahwa model semakin baik dalam meminimalkan kesalahan prediksi selama proses pelatihan. Namun, penurunan loss pada data validasi tidak sejalan dengan data pelatihan. Meskipun pada awalnya loss validasi mengalami penurunan, setelah mencapai titik tertentu, loss validasi mulai menunjukkan peningkatan atau stabil pada nilai yang lebih tinggi. Pola ini merupakan indikasi adanya overfitting, di mana model terlalu menyesuaikan diri dengan data pelatihan dan kehilangan kemampuan generalisasi pada data yang belum pernah dilihat sebelumnya.

Dalam grafik model Accuracy, terlihat akurasi pada data pelatihan terus meningkat seiring dengan bertambahnya epoch. Hal ini menunjukkan bahwa model menjadi semakin akurat dalam melakukan prediksi pada data yang digunakan untuk pelatihan. Namun, peningkatan akurasi pada data validasi tidak sebesar pada data pelatihan. Akurasi validasi cenderung stagnan setelah beberapa epoch, dan bahkan menunjukkan fluktuasi yang tidak sejalan dengan akurasi pelatihan, mengindikasikan bahwa model mengalami kesulitan dalam mempertahankan kinerja yang baik pada data baru yang tidak dilihat selama pelatihan.

D. Evaluasi Model Prediksi Hujan

Hasil evaluasi pada Tabel 2. menunjukkan performa dari tiga model yaitu CNN, RNN dan Transformer dalam melakukan tugas klasifikasi baik pada data pelatihan maupun data uji. Meskipun masing-masing model memiliki kekuatan dan kelemahan, hasil keseluruhan menunjukkan bahwa ketiga model ini belum memberikan performa yang memuaskan, dengan hasil uji dinilai "Kurang Baik".

CNN menunjukkan akurasi uji sebesar 56.85%, yang lebih baik daripada RNN (50.00%) tetapi lebih rendah dari Transformer (60.96%). Akurasi ini menunjukkan seberapa sering model melakukan prediksi yang benar pada data uji. Namun, akurasi saja tidak cukup untuk menilai performa model, terutama jika data yang digunakan memiliki distribusi kelas yang tidak seimbang. F1-Score CNN sebesar 45.22% menunjukkan bahwa meskipun model ini memiliki keseimbangan antara precision dan recall yang cukup baik, namun masih jauh dari optimal. Dengan precision sebesar 35.14% dan recall sebesar 63.41%, model CNN cenderung memberikan lebih banyak hasil positif yang benar (*high recall*) tetapi dengan banyak prediksi positif yang salah (*low precision*).

RNN memiliki performa yang paling rendah di antara ketiga model dengan akurasi uji 50.00%. Hal ini diikuti oleh F1-Score sebesar 36.52%, precision sebesar 28.38%, dan recall sebesar 51.22%. Kinerja yang rendah ini bisa disebabkan oleh kelemahan RNN dalam menangani urutan data yang panjang dan masalah gradien yang menghilang, yang sering kali sulit untuk menangkap hubungan jangka panjang. AUC RNN yang hanya mencapai 53.54% mengindikasikan bahwa kemampuan model dalam membedakan antara kelas positif dan negatif juga sangat terbatas. Selain itu korelasi antara prediksi dan target pada data pelatihan (0.13) dan data uji (0.06) menunjukkan bahwa model RNN

memiliki korelasi yang sangat lemah dengan target sebenarnya, yang memperlihatkan bahwa model tidak mampu menangkap pola yang signifikan dalam data.

Tabel 2. Evaluasi Model Prediksi Hujan

Metrik	CNN	RNN	Transformer
Akurasi (*Test*)	56.85%	50.00%	60.96%
F1-Score	45.22%	36.52%	50.00%
Precision	35.14%	28.38%	39.00%
Recall	63.41%	51.22%	71.00%
ROC AUC	67.34%	53.54%	68.00%
RMSE	0.49	0.5	0.49
MAE	0.47	0.5	0.46
Korelasi (*Train*)	0.37	0.13	0.38
Korelasi (*Test*)	0.27	0.06	0.29
Selisih Korelasi	-0.1	-0.07	-0.09
Rata-Rata Error (*Train*)	0.44	0.49	0.43
Akurasi (*Train*)	65.12%	53.75%	66.25%
Hasil Uji	Kurang Baik	Kurang Baik	Kurang Baik

Sementara itu model Transformer menunjukkan performa terbaik di antara ketiga model dengan akurasi uji sebesar 60.96%. F1-Score sebesar 50.00% menunjukkan bahwa Transformer berhasil mencapai keseimbangan antara precision (39.00%) dan recall (71.00%) yang lebih baik dibandingkan dua model lainnya. Tingginya recall menunjukkan bahwa Transformer mampu mengenali banyak instance dari kelas positif, meskipun precision yang relatif rendah menunjukkan bahwa banyak prediksi positif yang sebenarnya salah. AUC sebesar 68.00% juga lebih tinggi dibandingkan CNN dan RNN, menunjukkan bahwa Transformer lebih mampu dalam membedakan antara kelas-kelas yang ada. Dengan korelasi pada data pelatihan dan data uji masing-masing sebesar 0.38 dan 0.29, Transformer menunjukkan korelasi yang sedikit lebih kuat dengan target sebenarnya dibandingkan dua model lainnya.

Meski demikian, ketiga model menunjukkan perbedaan korelasi antara data pelatihan dan data uji yang negatif (selisih korelasi negatif), yang mengindikasikan adanya overfitting. Sebagaimana terlihat dari selisih korelasi yang negatif pada CNN (-0.1), RNN (-0.07), dan Transformer (-0.09). Rata-rata error (MAE dan RMSE) yang cukup serupa di ketiga model juga menunjukkan bahwa meskipun ada perbedaan dalam metrik performa, tingkat kesalahan prediksi pada skala absolut tidak terlalu bervariasi di antara model-model ini.

Hasil evaluasi ini juga mengungkap bahwa semua model menunjukkan hasil yang kurang baik pada data uji dibandingkan dengan data pelatihan, seperti yang ditunjukkan oleh penurunan akurasi dan peningkatan error. Hal ini menegaskan bahwa tidak satu pun dari ketiga model ini yang cukup kuat untuk generalisasi pada data uji, meskipun Transformer menunjukkan potensi yang lebih baik dibandingkan CNN dan RNN.

E. Model Prediksi Terbaik

Tabel 3. menampilkan dua kolom utama, satu kolom berisi nilai aktual apakah terjadi hujan atau tidak (hujan), dan kolom lainnya berisi prediksi dari model Transformer (prediksi (%). Jika kolom

hujan bernilai 0, prediksi yang ideal adalah mendekati 0%, dan jika bernilai 1, prediksi yang ideal adalah mendekati 100%. Prediksi yang mendekati nilai ideal ini menunjukkan kinerja model yang baik dalam mengklasifikasikan kondisi cuaca. Kolom lainnya merupakan data inputan untuk model prediksi terbaik untuk menghasilkan nilai prediksi (%) yang dinyatakan dalam bentuk persentase. Langkah pertama dalam menilai performa model ini adalah dengan menghitung akurasi, yaitu seberapa sering model ini berhasil memprediksi dengan benar apakah hujan akan terjadi atau tidak. Akurasi dihitung dengan cara membandingkan prediksi model dengan data aktual. Dari data yang diberikan, maka dihitung jumlah prediksi yang benar (baik prediksi 0 ketika tidak hujan maupun prediksi 1 ketika hujan) dan membaginya dengan total jumlah data. Nilai prediksi < 50% akan dianggap sebagai 0 dan untuk nilai prediksi > 50% akan dianggap sebagai 1 agar nilainya sama dengan nilai pada kolom hujan.

Pada Tabel 2 model membuat 850 prediksi. Setelah melakukan evaluasi, diperoleh hasil bahwa model berhasil memprediksi dengan benar sebanyak 433 kasus, sehingga akurasi model adalah:

$$Akurasi = \frac{433}{850}\ x\ 100\% = 50.94\%$$

Hasil akurasi 50.94% menunjukkan bahwa model hanya mampu memprediksi sekitar separuh dari data dengan benar, yang menunjukkan adanya ruang yang cukup signifikan untuk peningkatan kinerja model. Pada data prediksi hujan, terdapat beberapa kasus di mana nilai hujan adalah 1 (hujan terjadi), namun prediksi model tidak konsisten. Misalnya, ada beberapa kasus di mana nilai hujan adalah 1, tetapi prediksi menunjukkan nilai 0%, yang berarti model gagal mendeteksi adanya hujan. Kasus ini mengindikasikan bahwa model kesulitan dalam menangkap pola yang benar ketika hujan benar-benar terjadi. Sebaliknya, untuk kasus ketika nilai hujan adalah 0 (tidak hujan), terdapat juga beberapa kasus di mana prediksi model menunjukkan adanya kemungkinan hujan yang cukup tinggi (nilai prediksi mendekati 100%). Kasus seperti ini menunjukkan prediksi positif false, di mana model memperkirakan akan terjadi hujan namun tidak ada hujan yang sebenarnya terjadi.

Dalam model prediksi cuaca, dua jenis kesalahan yang paling signifikan adalah false positives (FP) dan false negatives (FN). FP terjadi ketika model memprediksi hujan akan terjadi, tetapi kenyataannya tidak (0-1), sementara FN terjadi ketika model tidak memprediksi hujan padahal kenyataannya hujan terjadi (1-0). Dengan data yang digunakan, terdapat beberapa kasus FP dan FN yang dapat menyebabkan model kurang dapat diandalkan dalam prediksi operasional cuaca dan perlu dilakukan peningkatan diberbagai sektor baik dari data pelatihan sampai dengan konfigurasi model.

Tabel 3. Matriks Kesalahan Prediksi Hujan Model Transformer

Kondisi Aktual	Prediksi Model	Jumlah	Keterangan
Hujan (1)	Hujan (1)	234	*True Positive (TP)*
Tidak Hujan (0)	Tidak Hujan (0)	199	*True Negative (TN)*
Hujan (1)	Tidak Hujan (0)	188	*False Negative (FN)*
Tidak Hujan (0)	Hujan (1)	229	*False Positive (FP)*
Total		**850**	

IV. SIMPULAN

Berdasarkan pembahasan yang telah dilakukan, hasil dari evaluasi model prediksi hujan menunjukkan bahwa model Transformer adalah yang paling efektif di antara ketiga model yang diuji yaitu CNN, RNN, dan Transformer. Dengan akurasi uji sebesar 60.96%, model prediksi Transformer menunjukkan kemampuan terbaik dalam mendeteksi pola cuaca kompleks dan memberikan prediksi yang lebih akurat dibandingkan dengan CNN dan RNN. Meskipun hasil prediksi masih menunjukkan

beberapa kesalahan, seperti false positives (FP) dan false negatives (FN), model ini memberikan kinerja yang lebih solid dalam hal akurasi, F1-Score, dan juga AUC. Model Transformer dapat menjadi salah satu opsi tambahan untuk memperkuat proses pengambilan keputusan bagi prakirawan BMKG yang bertugas untuk membuat prediksi cuaca di wilayah Panjang Bandar Lampung. Model Transformer memiliki potensi besar dalam sistem prediksi multi-model, memanfaatkan kekuatannya untuk meningkatkan prediksi cuaca secara keseluruhan. Untuk hasil yang optimal, disarankan untuk melanjutkan perbaikan dan penyesuaian pada model, termasuk eksplorasi lebih lanjut terhadap teknik optimisasi dan integrasi dengan model-model lain serta penggunaan data yang lebih variatif.

DAFTAR PUSTAKA

[1] A. Badrinath, L. D. Monache, N. Hayatbini, W. Chapman, F. Cannon, and M. Ralph, "Improving Precipitation Forecasts with Convolutional Neural Networks," *Weather Forecast.*, vol. 38, no. 2, pp. 291–306, 2023.

[2] R. Meka, A. Alaeddini, and K. Bhaganagar, "A robust deep learning framework for short-term wind power forecast of a full-scale wind farm using atmospheric variables," *Energy*, vol. 221, pp. 1–30, 2021.

[3] W. Wang, T. Ji, and J. Sun, "Prediction of Sea Clutter Based on Recurrent Neural Network," *J. Phys. Conf. Ser.*, vol. 1735, no. 1, pp. 1–17, 2021.

[4] D. Z. Haq *et al.*, "Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data," *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 829–837, 2021.

[5] I. Zahidin, B. Kanata, L. A. S. I. Akbar, "Perkiraan Suhu Menggunakan Algoritma Recurrent Neural Network Long Short Term Memory", *Journal of Computer System and Informatics* (JoSYC), vol.6, no.1. 244-252, 2024

[6] A. Vaswani *et al.*, "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.

[7] S. Ang, "Precipitation Forecasting Using Transformer: A Comparative Study with Unet," *Highlights Sci. Eng. Technol.*, vol. 39, pp. 627–632, 2023.

[8] I. K. Olarewaju and K. K. Kim, "Development of a Weather Prediction Device Using Transformer Models and IoT Techniques," *J. Sens. Sci. Technol.*, vol. 32, no. 3, pp. 164–168, 2023.

[9] J. Piran, X. Wang, H. Jun, and H. Han, "International Journal of Applied Earth Observation and Geoinformation Precipitation nowcasting using transformer-based generative models and transfer learning for improved disaster preparedness," Int. J. Appl. Earth Obs. Geoinf., vol. 132, no. June, p. 103962

[10] E. Bitsa, H. Flocas, J. Kouroutzoglou, M. Hatzaki, I. Rudeva, and I. Simmonds, "Development of a front identification scheme for compiling a cold front climatology of the Mediterranean," *Climate*, vol. 7, no. 11, pp. 9–13, 2019.

[11] J. Ko, K. Lee, H. Hwang, and K. Shin, "Deep-Learning-Based Precipitation Nowcasting with Ground Weather Station Data and Radar Data," *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2022-Novem, pp. 1063–1070, 2022.