ANALISIS POLA CUACA EXTREM MENGGUNAKAN *LONG SHORT TERM MEMORY* PADA WILAYAH JAWA BARAT

Muhammad Maulana Yusuf [1], Martanto [2], Umi Hayati [3]
[1,3] Teknik Informatika, [2] Manajemen Informatika, STMIK IKMI Cirebon
Jalan Perjuangan No.10B, Karyamulya, Kota Cirebon, Indonesia
Youseuf2802@gmail.com

ABSTRAK

Perubahan alam dipengaruhi oleh berbagai faktor terutama manusia, kegiatan ekonomi yang tidak mungkin terlepas dengan manusia, menjadi faktor utama. Perubahan lingkungan berdampak berubahnya cuaca, seperti pembangunan Gedung, *infrastruktur*, dan *deforestasi*, memengaruhi berubahnnya lingkungan, secara langsung mengakitbatkan berubahnya *vegetasi alam*, serta cuaca di wilayah tersebut ikut berubah. Perubahan cuaca yang tidak menentu, mengakibatkan munculnya Cuaca extrem, yaitu *fenomena* yang tidak biasa terjadi serta dapat menjadi masalah serius. Pada berbagai aspek kehidupan Kondisi seperti banjir dan badai contoh hasil cuaca extrem. Cuaca extrem mengakibatkan kerugian bagi mahluk hidup. penelitian ini bertujuan menganalisis terjadinya cuaca extrem menggunakan *Long Short Term Memory (LSTM). Model LSTM* digunakan untuk mengidentifikasi cuaca extrem, meramalkan terjadinya cuaca extrem di masa depan, memahami faktor yang memengaruhi munculnya cuaca extrem. Model LSTM mampu mengenali pola cuaca extrem dengan akurasi yang memadai. Penelitian ini mengembangkan model LSTM menggunakan 8 kali nilai treshlod value yaitu 28, 31, 34, 37, 40, 43, 46, 49, epoch sebanyak 50 dan model akurasi untuk menentukan keadaan cuaca mencapai 100 %. Data diambil dari *website BMKG* untuk digunakan pada penelitian. Dari hasil yang didapat, disimpulkan bahwa perbedaan nilai treshlod berpengaruh terhadap nilai prediksi cuaca kedapan selama 7 hari.

Kata kunci : *LSTM (Long Shor Term Memory), Extrem, Deforestasi, BMKG*

1. PENDAHULUAN

Teknologi mengalami perkembangan yang sangat pesat, munculnya pengetahuan baru menjadi ilmu yang bermanfaat, baik dibidang Pendidikan, ekonomi, cuaca, dan lainnya. Tujuan perkembangan dan perubahan, untuk memudahkan aktifitas manusia. Kecerdasan buatan muncul, membuat kumpulan data dapat diolah menjadi pengetahuan yang memiliki nilai manfaat. Kecerdasan buatan adalah sebuah system algoritma yang dibuat dan dikembangkan serta memiliki kemampuan berinovasi pada bidang yang dimodelkan baik komputer atau mesin, menjadi mampu berfikir seperti kecerdasan manusia atau bahkan melebihi[1].

Dengan data yang semakin banyak, akurasi juga meningkat. Cuaca adalah kondisi keadaan alam dalam rentang waktu yang terbatas dan sempit [2]. Pemahaman situasi cuaca, memiliki manfaat yang besar, karena memberi pengetahuan untuk pencegahan awal saat cuaca extrem. Perkembagan teknologi yang semakin modern memberi peluang meningkatkan akurasi ketepatan prediksi cuaca. Prediksi cuaca adalah Tindakan untuk memahami, bagaimana perkembangan keadaan udara dari masa lalu, saat ini, hingga memprediksi masa depan, untuk antisipasi apabila keadaan tidak normal [3].

Memahami analisis dan prediksi cuaca sangat penting, penelitian dalam bidang ini terus mengalami perkembangan. Akurasi yang didapatkan dengan menggunakan *esemble learning* sebesar 81.21% dengan *MSE* 18.79%, dengan metode *decision tree algoritma forest* akurasi yang didapat sebesar 82.38% dengan *MSE* 17.62%, dengan metode *deep learning* mendapat akurasi sebesar 82.92% dengan *MSE* 17.08%. *MSE (Mean Squared Error)* adalah salah satu metrik, yang digunakan untuk prediksi cara dan mengukur sejauh mana model cocok dengan data pengamatan aktual, serta prediksi cuaca sejauh mana perkiraan model akurat [3]. Selain itu, proses pemantauan integrasi radar cuaca diseluruh wilayah Indonesia secara Bersama, ternyata masih memperlukan waktu yang cukup lama [4].

Memprediksi cuaca extrem, menjadi tantangan bagi para ahli pemodelan di Indonesia bahkan dunia. Cuaca extrem menjadi permasalah kompleks karena peluang terjadinya kecil mengakibatkan model yang dikembangkan memiliki akurasi yang rendah [5]. Tujuan utama penelitian ini adalah untuk meningkatkan akurasi yang lebih tinggi, dalam memprediksi dan menganalisis cuaca extrem yang lebih akurat untuk masa mendatang.

Penelitian ini menggunakan metode *deep learning* yaitu algoritma *Long Short Term Memory (LSTM)*, mengembangkan lebih lanjut berdasarkan data. *LSTM* merupakan metode jaringan syaraf tiruan, yang menggunakan informasi dari masa lalu, untuk diproses dengan bentuk data berurutan atau sekuensial [6]. Dengan menggunakan algoritma *LSTM*, mampu meningkatkan akurasi analisis cuaca, akan mendukung berbagai sektor untuk mengambil keputusan yang lebih tepat. Akurasi prediksi cuaca extrem yang tinggi, dapat memberikan informasi akurat, kemudian mengurangi risiko dampak cuaca extrem. Penggunaan teknologi *artificial intelegent* seperti *deep learning*, untuk analisis dan prediksi cuaca, memberikan kemungkinan perencanaan yang lebih baik. hasil

penelitian ini berpotensi besar untuk meningkatkan pemahaman tentang informasi cuaca, serta konstribusi teknologi ar*tificial intelegent metode deep learning algoritma LSTM,* untuk analisis dan memprediksi di bidang cuaca yang lebih baik.

2. TINJAUAN PUSTAKA

Pada penelitian ini metode yang digunakan adalah *systematic literature review. Metode systematic literature review* adalah metode dengan cara melakukan tela'ah dan mengolah sumber tulisan ilmiah yang ada serta telah terpublikasi[7].

2.1. Metode

Pada penelitian ini metode tahapan – tahapan tersebut adalah

a. Tahapan identifikasi, melakukan pencarian literatur sesuai dengan judul.
b. Tahapan seleksi, mealukan pemilihan dan penelompokan sesuai penelitian yang dilakukan.
c. Tahapan verifikasi, pemeriksaan berdasarkan kriteria yang ditetapkan peneliti.
d. Tahapan kombinasi dan menyimpulkan, hasil memadukan teori, yang telah ada dengan penelitian ini.

2.2. Hasil

Hasil yang telah dilakukan pada jurnal – jurnal terkait dengan judul penelitian.

a. Jurnal berjudul "analisis data time series menggunakan *LSTM (Long Short Term Memory) dan ARIMA (Autocorrelation Integrated Moving Average)* dalam *Bahasa Python*" menjelaskan, metode *statistic ARIMA (Autocorrelation Integrated Moving Average*) memerlukan peningkatan akurasi serta perlu menurunkan nilai *Error*. Penelitian ini juga, menggunakan metode *LSTM (Long Short Term Memory)* mendapat kesimpulan metode *LSTM* lebih akurat, dibandingkan menggunakan *ARIMA*. Hasil prediksi penelitian, berupa *saham* Telkom mencapai tingkat akurasi 99%, *RMSE* 1 %, setelah data melalui proses *transformasi loging* kemudian *EWMA* atau *loging* kemudian *Moving average*[8].
b. Jurnal berjudul "*implementasi deep learning* menggunakan *arsitektur Long Short Term Memory* untuk prediksi curah hujan kota malang" menjelaskan *deep learning*, yang dilakukan penelitian sebelumnya mencapai akurasi sebesar 85%. *Deep learning* merupakan jaringan, yang memiliki beberapa *layer* dimana *layer – layer* tersebut, berasal dari kumpulan *node – node.* Dengan menggunakan *arsitektur LSTM (Long Short Term Memory)* dalam kasus yang berbeda, *RME* dihasilkan kecil, kesimpulannya metode *LSTM (Long Short Term Memory)* untuk prediksi curah hujan kota malang, memberikan hasil prediksi yang tergolong cukup baik. Oleh karena itu arsitektur *LSTM* mampu bekerja cukup optimal[9] .
c. Jurnal berjudul "*implementasi Long Short Term Memory (LSTM)* untuk perdiksi intensitas curah hujan (studi kasus : kabupaten malang)" menjelaskan, kabupaten malang adalah daerah yang memiliki *iklim tropis* dan memiliki sumber daya alam yang melimpah, seperti sektor pertanian dan perkebunan, sektor ini memiliki beberapa faktor yang dapat mempengaruhi tingkat produktifitas, salah satunya yaitu curah hujan. Penelitian membahas prediksi curah hujan, di kabupaten malang dengan menggunakan metode LSTM (*Long Short Term Memory)*. Kesimpulannya metode *LSTM (Long Short Term Memory)* memiliki performa terbaik, dengan *parameter* yang ditentukan dengan nilai *error* adalah *RMSE* sebesar 0.98162 dan *MAE* sebesar 0.68847[10].

2.3. Tabel

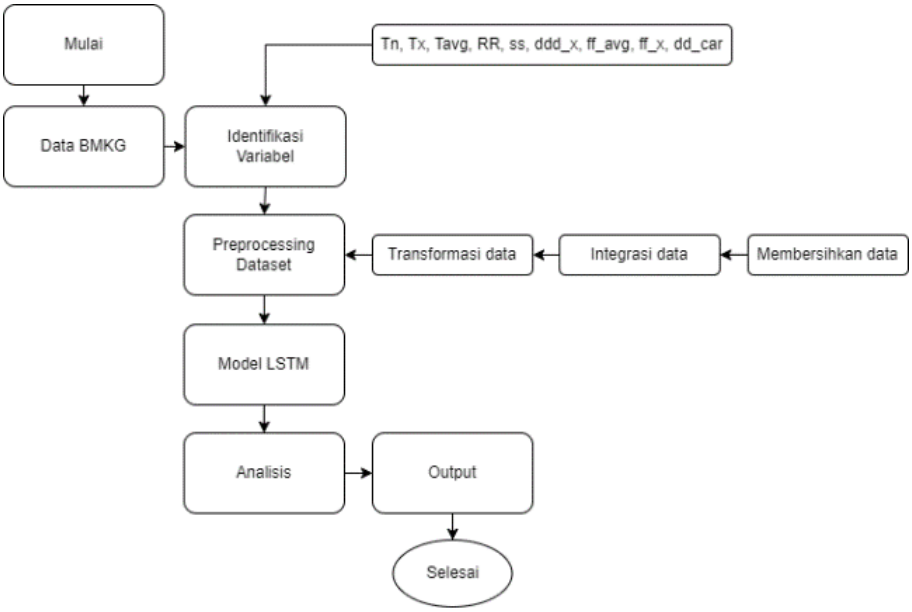
Tabel 1. Ringkasan

Judul	Author	Masalah	Metode
Analisis data *time series* menggunakan *lstm (long short term memory) dan Arima (autocorrelation integrated moving average) dalam bahasa python*	Adhitio satyo, bayangkari karno	Memprediksi *saham* Telkom.	*LSTM (Long Short Term Memory) dan ARIMA (Autocorrelation Integrated Moving Average).*
Implementasi deep learning menggunakan arsitektur *long Short-term memory* untuk prediksi curah hujan kota malang	Muhammad Rizki, Setio Basuki, Yufis Azhar	Memprediksi curah hujan di kota malang.	*Lstm (long short term memory).*
Implementasi long short-term memory (lstm) untuk prediksi intensitas curah hujan (studi kasus: kabupaten malang)	Thoriq Afa Faisal Muhammad, dan Mohammad Isa Irawan	Memprediksi *parameter cuaca*, yang berpengaruh terhadap lingkungan dimana kabupaten malang adalah daerah yang memiliki iklim tropis dan bergerak pada sektor pertanian dan perkebunan.	*Lstm (long short term memory).*

Berdasarkan tabel 1. Ringkasan, disimpulkan cuaca sangat mempengaruhi kegiatan manusia, baik itu disektor pertanian, perkebunan, transportasi, bisnis, atau bahkan aktivitas setiap individu. Cuaca extrem dapat membuat banyak kerugian seperti yang dijelaskan oleh jurnal yang berjudul "*tipologi* kerusakan bangunan akibat cuaca extrem studi kasus cuaca extrem kabupaten bantul 2019" menjelaskan, telah terjadi peristiwa bencana sebanyak 2.564 diseluruh wilaya Indonesia, yang menyebabkan banyak kerugian seperti rusaknya bangunan. Oleh karena itu, banyak penelitian yang mencoba untuk memprediksi keadaan cuaca dan menganalisis faktor -faktor terjadi cuaca extrem. Salah satu metode, yang mampu untuk menganalisis dan memprediksi cuaca adalah *algoritma LSTM (Long Short Term Memory)*. Tipe data yang digunakan adalah data *time series* atau urutan waktu, dengan cara melatih data tersebut menggunakan *epoch dan batch size*.

3. METODE PENELITIAN

Penelitian ini, masuk dalam kategori penelitian *kuantitatif*, menggunakan algortima *Long Short-Term Memory (LSTM)*. *LSTM* merupakan jenis arsitektur jaringan saraf tiruan, yang dirancang khusus untuk menangani pengolahan data urutan waktu.



Gambar 1. metode penelitian

Tabel 1. Metode penelitian

Tahap	Aktifitas	Deskripsi Aktifitas
1.pengambilan data dari *BMKG*	Download data, dari *website BMKG*	*Mendowload* data dari tanggal 1 januari 2019 sampai 31 desember 2022.
2.*Analisis Variabel*	Melakukan *analisis variabel*	*Analisis variabel*, yang akan digunakan, mendukung penelitian dengan *google colaboratory*.
3.*Preprocessing Data*	Melakukan *Preprocessing*, yang telah di *analisis*.	Melakukan *preprocessing* agar data sesuai model, yang digunakan dan menentukan sampel
4.Pembuatan *Model LSTM (Long Short – Term Memory)*.	Menerapkan model *LSTM*.	Menerpakan model *LSTM*.
5.*Analisis dan Prediksi*	Melakukan *analisis dan prediksi* hasil yang didapat	Melakukan *analisis*, yang didapat dan memprediksi nilai cuaca kedepan.

3.1. Sumber Data

Data cuaca diambil dari *website BMKG*, stasiun klimatologi jawa barat dengan lintang – 6.50000, bujur 106.75000 dan elevasi 207. Berikut alamat website (https://dataonline.bmkg.go.id/home). Pengambilan sampel dilakukan, dengan *mengidentifikasi* data cuaca yang memenuhi *kriteria* serta ditelah ditetapkan dengan *google colaboratori*.

3.2. Populasi dan Sampel

Populasi data mencakup semua rekaman data cuaca harian, dari tanggal 01/01/2019 sampai 31/12/2022 yang mencapai 1640 data. Sampel data yaitu *Tn (Temperatur minimum), Tx (Temperatur maksimum), Tavg (Temperatur rata - rata), RH_avg (kelembaban rata - rata), RR (Curah Hujan), ss (lamanya penyinaran matahari), ff_x (kecepatan angin maksimum), ddd_x (arah angin saat kecepatan maksimum), ff_aavg (kecepatan angin rata - rata), ddd_car (arah angin terbanyak).*

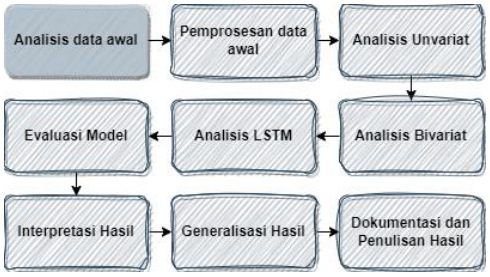
3.3. Teknik Pengambilan Data

Pengumpulan data secara online, yang diambil dari website resmi dari *BMKG (Badan Meteorologi, Klimatologi, dan Geofisika)* yang dapat di gunakan

secara legal, dengan mendownload data tersebut secara berkala, per satu bulan dari tanggal 1 januari 2019 sampai tanggal 31 desember 2022. Kemudian data di kumpulkan menjadi 1 file, yang berbentuk excel.

3.4. Teknik Analisis Data

Berikut merupakan gambaran dan penjelasan, dari Teknik analisis data yang akan dilakukan penelitian ini :



Gambar 2. teknik analisis data

Berikut Teknik analisis data penelitian ini :

a. *Analisis* data awal : untuk memahami kaarektersitik data cuaca.
b. Pemprosesan data : memastikan data dalam kondisi baik.
c. Analisis *unvariat* : memahami distribusi antar variabel, termasuk pola dan tren dalam data.
d. Analisis *bivariat* : memahami korelasi antar variabel.
e. Analisis *LSTM* : melatih model dengan algortima LSTM.
f. Evaluasi model : mengevaluasi mencakup penggunakan metrik seperti *MSE*.
g. Intepretasi hasil : memprediksi cuaca kedepan dengan menggunakan model *LSTM*.
h. Generalisasi hasil : meninjau Kembali hasil sampel data cuaca ke populasi secara keseluruan.
i. Dokumentasi dan penulisan hasil : membuat kesimpulan dan mendokumentasikan hasil yang telah didapat.

4. HASIL DAN PEMBAHASAN

Dibawah ini uraian hasil dari implementasi yang dilakukan sesuai dengan tahap metode penelitian.

4.1. Pengambilan data BMKG

Data diambil melalui website resmi *BMKG* yang dapat digunakan secara *legal*, berkut alamat websitenya (https://dataonline.bmkg.go.id/home). Dengan cara mendowloat per satu bulan.

4.2. Analisis variabel

Berikut tahapan Untuk menganalisis variabel yang tidak dibutuhkan pada tahap selanjutnya.

a. *Install Library*

library yang harus di *install* untuk digunakan, berikut library yang harus di *install.*

```
# Install library yang diperlukan
!pip install tensorflow numpy pandas matplotlib
```

Gambar 3. kode instal librari 1

PIP untuk *install tensorflow*, mengembangkan *machine learning*, *numpy*, *pandas*, dan *matplotlib* untuk *manipulasi*, *analisis*, dan *visulasasi data*.

b. *Import Library*

```
# Import library

import pandas as pd
import numpy as np
from numpy import array
from numpy import hstack
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler, StandardScaler
from keras.preprocessing.sequence import TimeseriesGenerator
import tensorflow as tf
from tensorflow.keras import optimizers
from keras.models import Sequential
from keras.layers import LSTM
from keras.layers import Dense
from keras.layers import RepeatVector
from keras.layers import TimeDistributed
from keras.layers import Dropout
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
from tensorflow.keras.callbacks import EarlyStopping
```

Gambar 4.kode install library 2

Import library yang dibutuhkan pengolahan data sesuai fungsinya, Seperti *pandas*, *numpy*, *matplotlib*, *seaborn*, dan lain sebagainya.

c. Mengakses Data

```
[ ] # import google drive
    from google.colab import drive
    drive.mount('/content/drive')

    Mounted at /content/drive

[ ] # import file
    data = pd.read_excel('/content/drive/MyDrive/6. TUGAS SEMESTER 8/cuaca/cuaca.xlsx')

    # merubah data excel menjadi data frame
    df = pd.DataFrame(data)

    print(df)

          Tanggal    Tn    Tx  Tavg  RH_avg    RR    ss  ff_x  ddd_x  ff_avg  \
0      01-01-2019  22.5  30.4  25.0    89.0   9.3   1.9  10.0  286.0     2.0
1      02-01-2019  23.4  32.1  26.0    84.0   1.6   1.8   5.0  260.0     1.0
2      03-01-2019  22.9  32.2  26.8    76.0   2.0   6.8   3.0  240.0     1.0
3      04-01-2019  21.3  32.8  26.6    77.0   NaN   5.6   4.0  320.0     1.0
4      05-01-2019  22.0  33.2  26.7    76.0   NaN  11.2   7.0  360.0     1.0
...           ...   ...   ...   ...     ...   ...   ...   ...    ...     ...
1455   27-12-2022  20.8  27.9  24.7    85.0  22.9   0.0   5.0  290.0     1.0
```

Gambar 5. Mengakses data

Untuk menghubungkan colaboratory dengan *google drive* agar dapat memanggil data yang akan digunakan.

d. *Analisis Varibel*

Beberapa variabel yang tidak dibutuhkan dalam penelitian ini adalah *Tn*, *Tx*, *ddd_x*, *ff_avg*, *ddd_car*. Berikut kode yang digunakan untuk menghapus variabel yang tidak dibutuhkan.

```
# Menghapus kolom-kolom tertentu dari DataFrame
hapus_data = ['Tn', 'Tx', 'ddd_x', 'ff_avg', 'ddd_car']
df = df.drop(hapus_data, axis=1)

# Menampilkan DataFrame setelah menghapus kolom
print(df)

          Tanggal  Tavg  RH_avg   RR   ss  ff_x
0      01-01-2019  25.0    89.0  9.3  1.9  10.0
1      02-01-2019  26.0    84.0  1.6  1.8   5.0
2      03-01-2019  26.8    76.0  2.0  6.8   3.0
3      04-01-2019  26.6    77.0  NaN  5.6   4.0
4      05-01-2019  26.7    76.0  NaN 11.2   7.0
...           ...   ...     ...  ...  ...   ...
1455   27-12-2022  24.7    85.0 22.9  0.0   5.0
1456   28-12-2022  25.5    82.0  1.0  0.0   4.0
1457   29-12-2022  24.8    92.0 36.7  0.1   5.0
1458   30-12-2022  24.4    92.0 18.7  0.0   3.0
1459   31-12-2022  24.3    92.0  1.5  0.0  12.0

[1460 rows x 6 columns]
```

Gambar 6. Kode menghapus variabel dan membaca data

variabel yang memiliki nilai *NaN*, *null*, *9999*, *8888*. Dalam penelitian ini, data *missing value* tersebut akan di atasi dengan *Teknik interpolate.*

```
[7] # Mengganti nilai-nilai NaN, null, 9999, atau 8888 dengan 0 di seluruh DataFrame
    df = df.fillna(np.NaN)  # Ganti NaN dengan 0
    df = df.replace([9999, 8888, None], np.NaN)  # Ganti 9999, 8888, dan null dengan 0

[8] # Interpolasi linear untuk mengisi nilai yang hilang
    df.interpolate(method='linear', inplace=True)

    # Menampilkan DataFrame setelah mengganti nilai
    print(df)

         Tanggal  Tavg  RH_avg    RR   ss  ff_x
0     01-01-2019  25.0    89.0  9.30  1.9  10.0
1     02-01-2019  26.0    84.0  1.60  1.0   5.0
2     03-01-2019  26.0    76.0  2.00  6.0   3.0
3     04-01-2019  26.0    77.0  4.26  5.5   4.0
4     05-01-2019  26.7    76.0  6.52 11.2   7.0
...          ...   ...     ...   ...  ...   ...
1455  27-12-2022  24.7    85.0 22.90  0.0   5.0
1456  28-12-2022  23.6    93.0  5.00  0.0   4.0
1457  29-12-2022  24.0    92.0 26.20  0.1   3.0
1458  30-12-2022  24.4    92.0 18.70  0.0   3.0
1459  31-12-2022  24.3    92.0  1.50  0.0  12.0

[1460 rows x 6 columns]
```

Gambar 7. kode teknik interpolate

4.3. Preprocessing data

Hal pertama adalah memanggil data yang telah dibersihkan. Berikut kode yang digunakan, untuk memanggil data dan membacanya.

```
[10] # import file yang telah dibersihkan
     data = pd.read_excel('/content/drive/MyDrive/6. TUGAS SEMESTER 8/cuaca/cuaca_clear.xlsx')

     # membaca data
     print(data)

         Tanggal  Tavg  RH_avg    RR   ss  ff_x
0     01-01-2019  25.6    89.0  9.30  1.9  10.0
1     02-01-2019  26.0    84.0  1.60  1.0   5.0
2     03-01-2019  26.0    76.0  2.00  6.0   3.0
3     04-01-2019  26.6    77.0  4.26  5.5   4.0
4     05-01-2019  26.7    76.0  6.52 11.2   7.0
...          ...   ...     ...   ...  ...   ...
1455  27-12-2022  24.7    85.0 22.90  0.0   5.0
1456  28-12-2022  23.6    93.0  5.00  0.0   4.0
1457  29-12-2022  24.0    92.0 26.20  0.1   3.0
1458  30-12-2022  24.4    92.0 18.70  0.0   3.0
1459  31-12-2022  24.3    92.0  1.50  0.0  12.0

[1460 rows x 6 columns]
```

Gambar 8. Kode import data bersih 1

Data *bertipe float* perlu dirubah menjadi *integer* agar mempermudah menentukan kondisi cuaca. Berikut kode yang digunakan.

```
# Daftar atribut yang ingin Anda ubah menjadi integer
atribut_float_to_int = ['Tavg', 'RH_avg', 'RR', 'ss', 'ff_x']

# Menggunakan metode astype() untuk mengubah tipe data menjadi integer
data[atribut_float_to_int] = data[atribut_float_to_int].astype(int)

# Tampilkan beberapa baris pertama data setelah perubahan
data.head()

     Tanggal  Tavg  RH_avg  RR  ss  ff_x
0  01-01-2018    25      89   9   1    10
1  02-01-2018    26      84   1   1     5
2  03-01-2018    26      76   2   6     3
3  04-01-2018    26      77   4   5     4
4  05-01-2018    26      76   6  11     7
```

Gambar 9. kode merubah data *float* menjadi *integer*

Kemudian memberikan label kondisi sesuai cuaca yang dapat terjadi di Indonesia, dengan membandingkan nilai *variabel RR (Curah Hujan) dan ss (Intensitas Matahari), menggunakan teknik if else.* Berikut kode yang digunakan.

```
# Fungsi untuk menentukan kondisi cuaca
def tentukan_kondisi_cuaca(row):
    if (
        row['RR'] > row['ss']
    ):
        return "HUJAN"
    elif (
        row['RR'] < row['ss']
    ):
        return "TERANG"
    elif (
        row['RR'] == row['ss']
    ):
        return "MENDUNG"

    # Default jika tidak memenuhi kondisi di atas
    return "Tidak Diketahui"

# Menambahkan kolom 'Kondisi_Cuaca' ke dalam DataFrame
data['Kondisi_Cuaca'] = data.apply(tentukan_kondisi_cuaca, axis=1)

# Menampilkan DataFrame dengan kolom 'Kondisi_Cuaca'
print("\nData cuaca dengan kondisi:")
print(data[['Tanggal', 'Tavg', 'RH_avg', 'RR', 'ss', 'ff_x', 'Kondisi_Cuaca']])
```

Gambar 10. Kode menentukan kondisi cuaca

Kondisi cuaca yaitu hujan, mendung, dan terang. Terjadinya hujan apabila nilai *RR (Curah Hujan) lebih besar dari nilai ss (Intesitas Matahari), mendung apabila nilai RR (Curah Hujan) sama dengan nilai ss (Intesitas Matahari), dan terang apabila nilai RR (Curah Hujan) lebih kecil dari nilai ss (Intesitas Matahari).* Berikut hasil dari kode diatas.

```
Data cuaca dengan kondisi:
         Tanggal  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca
0     01-01-2019    25      89   9   1    10         HUJAN
1     02-01-2019    26      84   1   1     5       MENDUNG
2     03-01-2019    26      76   2   6     3        TERANG
3     04-01-2019    26      77   4   5     4        TERANG
4     05-01-2019    26      76   6  11     7        TERANG
...          ...   ...     ...  ..  ..   ...           ...
1455  27-12-2022    24      85  22   0     5         HUJAN
1456  28-12-2022    23      92   3   0     4         HUJAN
1457  29-12-2022    24      92  26   0     3         HUJAN
1458  30-12-2022    24      92  18   0     3         HUJAN
1459  31-12-2022    24      92   1   0    12         HUJAN

[1460 rows x 7 columns]
```

Gambar 11. hasil kondisi cuaca

Menentukan terjadinya cuaca extrem, dengan menghitung nilai rata – rata setiap kolom. Berikut kode untuk membuat rata – rata setiap kolom dan hasilnya.

```
# Membuat kolom baru 'Rata_rata' yang berisi rata-rata dari kolom yang diinginkan
data['Rata_rata'] = data[['Tavg', 'RH_avg', 'RR', 'ss', 'ff_x']].mean(axis=1)

# Menampilkan DataFrame yang telah diperbarui
print(data)

         Tanggal  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca  Rata_rata
0     01-01-2019    25      89   9   1    10         HUJAN       26.8
1     02-01-2019    26      84   1   1     5       MENDUNG       23.4
2     03-01-2019    26      76   2   6     3        TERANG       22.6
3     04-01-2019    26      77   4   5     4        TERANG       23.2
4     05-01-2019    26      76   6  11     7        TERANG       25.2
...          ...   ...     ...  ..  ..   ...           ...        ...
1455  27-12-2022    24      85  22   0     5         HUJAN       27.2
1456  28-12-2022    23      92   3   0     4         HUJAN       24.4
1457  29-12-2022    24      92  26   0     3         HUJAN       29.0
1458  30-12-2022    24      92  18   0     3         HUJAN       27.4
1459  31-12-2022    24      92   1   0    12         HUJAN       25.8

[1460 rows x 8 columns]
```

Gambar 12. kode menentukan nilai rata – rata

Setelah memberikan nilai rata – rata dapat dilihat nilai tertinggi, nilai terendah, dan nilai rata – rata, dari kolom rata-rata, yang digunakan sebagai batas terjadinya cuaca ekstrem. Berikut kode yang digunakan dan hasilnya.

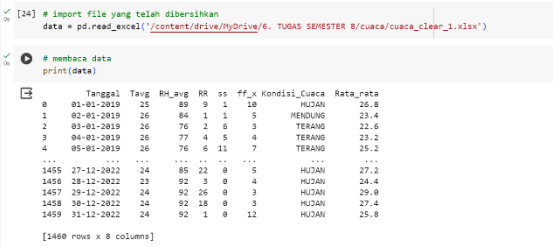
```
# Cari rata-rata tertinggi
rata_rata_tertinggi = data['Rata_rata'].max()

# Tampilkan hasil
print("nilai (Rata-rata) Tertinggi:", rata_rata_tertinggi)

# Cari rata-rata terendah
rata_rata_terendah = data['Rata_rata'].min()

# Tampilkan hasil
print("nilai (Rata-rata) Terendah:", rata_rata_terendah)

# Cari rata-rata dari nilai rata-rata
rata_rata_dari_rata_rata = data['Rata_rata'].mean()

# Tampilkan hasil
print("nilai (Rata-rata) dari nilai rata-rata:", rata_rata_dari_rata_rata)

nilai (Rata-rata) Tertinggi: 56.2
nilai (Rata-rata) Terendah: 20.6
nilai (Rata-rata) dari nilai rata-rata: 26.846986301369867
```

Gambar 13. kode menentukan nilai

Didapat nilai tertinggi adalah 56.2, nilai terendah 20.6, dan rata – rata 26.84. maka nilai cuaca normal di jawa barat sesuai dengan data adalah 26.84 sebagai kondisi nilai cuaca rata – rata harian.

4.4. Pembuatan model Long Short Term Memori

Selanjutnya membuat model *Long Short Term Memory. Import* Kembali data yang telah di proses. Berikut kode yang digunakan untuk memanggi data dan menampilkannya.

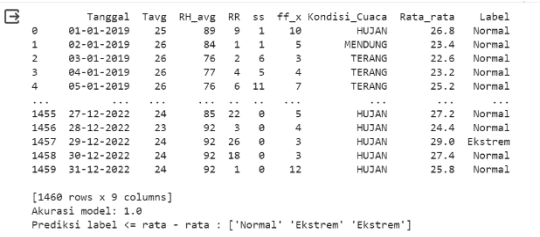


```
[24] # import file yang telah dibersihkan
     data = pd.read_excel('/content/drive/MyDrive/6. TUGAS SEMESTER 8/cuaca/cuaca_clear_1.xlsx')

     # membaca data
     print(data)

          Tanggal  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca  Rata_rata
0     01-01-2019    25      89   9   1    10         HUJAN       26.8
1     02-01-2019    26      84   1   1     5      MENDUNG       23.4
2     03-01-2019    26      76   2   6     3       TERANG       22.6
3     04-01-2019    26      77   4   5     4       TERANG       23.2
4     05-01-2019    26      76   6  11     7       TERANG       25.2
...          ...   ...     ...  ..  ..   ...          ...        ...
1455  27-12-2022    24      85  22   0     5        HUJAN       27.2
1456  28-12-2022    23      92   3   0     4        HUJAN       24.4
1457  29-12-2022    24      92  26   0     3        HUJAN       29.0
1458  30-12-2022    24      92  18   0     3        HUJAN       27.4
1459  31-12-2022    24      92   1   0    12        HUJAN       25.8

[1460 rows x 8 columns]
```

Gambar 14. kode membaca data bersih 2

Kemudian menentukan terjadinya cuaca extrem atau cuaca normal, dari nilai rata – rata yang di dapat yaitu 26.84. semakin tinggi nilai rata – rata setiap kolom, kondisi yang terjadi semakin extrem dan semakin jarang terjadi. Berikut kode yang digunakan untuk menentukan kondisi cuaca.

Gambar 15. Kode Label dan akurasi data



```
          Tanggal  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca  Rata_rata     Label
0     01-01-2019    25      89   9   1    10         HUJAN       26.8    Normal
1     02-01-2019    26      84   1   1     5      MENDUNG       23.4    Normal
2     03-01-2019    26      76   2   6     3       TERANG       22.6    Normal
3     04-01-2019    26      77   4   5     4       TERANG       23.2    Normal
4     05-01-2019    26      76   6  11     7       TERANG       25.2    Normal
...          ...   ...     ...  ..  ..   ...          ...        ...       ...
1455  27-12-2022    24      85  22   0     5        HUJAN       27.2    Normal
1456  28-12-2022    23      92   3   0     4        HUJAN       24.4    Normal
1457  29-12-2022    24      92  26   0     3        HUJAN       29.0  Ekstrem
1458  30-12-2022    24      92  18   0     3        HUJAN       27.4    Normal
1459  31-12-2022    24      92   1   0    12        HUJAN       25.8    Normal

[1460 rows x 9 columns]
Akurasi model: 1.0
Prediksi label <= rata - rata : ['Normal' 'Ekstrem' 'Ekstrem']
```

Gambar 16. Hasil dari kode label

Didapat akurasi dengan model *RandomforestClassifier*, sebagai model untuk melihat *akurasi* dalam memberikan label dengan state 42, didapat akurasi 1.0 yang memiliki arti 100% akurat dalam menentukan kondisi normal atau extrem. Selanjutnya menerapkan model *algoritma Long Short Term Memory* dan melatih model epoch sebanyak 50 kali dengan berbagai nilai *treshlod*, dan nilai 26.84 yang dapat dibulatkan 27 sebagai acuan nilai rata – rata kondisi normal cuaca harian di jawa barat, serta memprediksi nilai cuaca yang terjadi selama 7 hari kedepan. Berikut kode

Tabel 2. Kode Long Short Term Memory

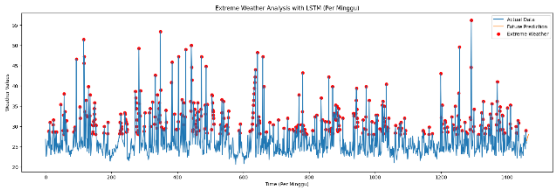
```
# memanggil data yang telah di bersihkan
file_path = '/content/drive/MyDrive/6. TUGAS SEMESTER 8/cuaca/cuaca_clear_1.xlsx'
# Membaca dataset cuaca yang telah dibersihkan
df = pd.read_excel(file_path)
# Menentukan batas nilai 'Rata_rata' untuk diwarnai
threshold_value = (memberikan nilai ke ekstriman yang terjadi dengan dari 27 - 56)
# Menambahkan kolom baru 'Extreme' sebagai penanda cuaca ekstrem
df['Extreme'] = np.where(df['Rata_rata'] >= threshold_value, 1, 0)
# Menghitung nilai rata-rata dan menormalisasi data
average_value = df['Rata_rata'].mean()
scaler = MinMaxScaler(feature_range=(0, 1))
scaled_data = scaler.fit_transform(df['Rata_rata'].values.reshape(-1, 1))
# Membagi data menjadi set pelatihan dan pengujian
train_size = int(len(scaled_data) * 0.80)
test_size = len(scaled_data) - train_size
train, test = scaled_data[0:train_size, :], scaled_data[train_size:len(scaled_data), :]
# Membuat dataset untuk LSTM
def create_dataset(dataset, time_step=1):
    dataX, dataY = [], []
    for i in range(len(dataset) - time_step):
        a = dataset[i:(i + time_step), 0]
        dataX.append(a)
        dataY.append(dataset[i + time_step, 0])
    return np.array(dataX), np.array(dataY)
# Mengubah nilai time_step agar sesuai dengan kebutuhan prediksi di masa depan
time_step = 7  # menyesuaikan sesuai yang ingin di predikasi perhari/perminggu/perbulan
X_train, y_train = create_dataset(train, time_step)
X_test, y_test = create_dataset(test, time_step)
# Reshape data agar sesuai dengan input LSTM
X_train = X_train.reshape(X_train.shape[0], X_train.shape[1], 1)
X_test = X_test.reshape(X_test.shape[0], X_test.shape[1], 1)
# Membuat model LSTM
model = Sequential()
model.add(LSTM(units=50, return_sequences=True, input_shape=(X_train.shape[1], 1)))
model.add(LSTM(units=50, return_sequences=True))
model.add(LSTM(units=50))
model.add(Dense(units=1))
model.compile(optimizer='adam', loss='mean_squared_error')
# Melatih model
model.fit(X_train, y_train, epochs=50, batch_size=64, verbose=1)
```

4.5. Analisis dan prediksi

Berdasarkan nilai dari proses yaitu 26.84 dibulatkan menjadi 27 sebagai tolak ukur nilai cuaca normal harian pada wilayah jawa barat. berikut grafik kondisi cuaca.

Nilai *treshlod value* 28

Didapat grafik di bawah ini, dengan pola titik merah muncul cukup banyak. Terjadi sebanyak 392 kondisi di atas nilai *treshlod* value 28.

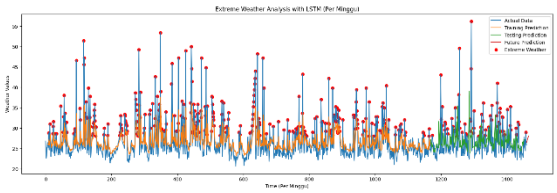


Gambar 17. grafik treshlod 28

```
Threshold Values:
          Tanggal  Rata_rata  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca    Label
9     10-01-2019       28.8    26      86  23   5     4         Hujan  Ekstrem
13    14-01-2019       31.0    25      90  34   4     2         Hujan  Ekstrem
22    23-01-2019       28.6    24      92  24   0     3         Hujan  Ekstrem
23    24-01-2019       30.4    24      91  32   2     3         Hujan  Ekstrem
24    25-01-2019       31.6    24      90  39   2     3         Hujan  Ekstrem
...          ...        ...   ...     ...  ..  ..   ...           ...      ...
1427  29-11-2022       30.0    26      82  32   6     4         Hujan  Ekstrem
1428  30-11-2022       28.4    26      84  21   7     4         Hujan  Ekstrem
1433  05-12-2022       31.4    25      88  35   5     4         Hujan  Ekstrem
1435  07-12-2022       30.8    25      88  33   3     5         Hujan  Ekstrem
1457  29-12-2022       29.0    24      92  26   0     3         Hujan  Ekstrem

[392 rows x 9 columns]
```

Gambar 18. hasil *treshlod* 28

Prediksi kondisi cuaca selama 7 hari kedapan menggunakan nilai *treshlod value* 28, didapat kondisi 7 hari kedepan seperti gambar dibawah ini.



Gambar 19. grafik prediksi treshlod 28

Time Step	Prediction	Condition	Weather
1460	[26.496822]	Normal	Terang
1461	[26.773525]	Normal	Terang
1462	[27.101135]	Normal	Terang
1463	[27.401043]	Normal	Terang
1464	[27.65401]	Normal	Terang
1465	[27.860886]	Normal	Terang
1466	[28.03082]	Extreme	Hujan

Gambar 20.hasil prediksi dengan treshlod 28

Selama 6 hari kedepan prediksi cuaca normal kondisi terang dan hujan dihari ke 7 dengan cuaca ekstrem. Nilai *treshlod value* dapat diinputkan sesuai keinginan dengan standar nilai cuaca normal di jawa barat adalah 26.84 yang dibulatkan menjadi 27. Grafik warna biru adalah nilai data *actual*, grafik warna oren adalah grafik data latih, grafik warna hijau adalah grafik data uji, dan grafik warna merah di bagian paling akhir adalah grafik prediksi sebenarnya.

Dari berbagai percobaan nilai *treshlod* diatas, semakin tinggi nilai *treshlod* yang diinputkan maka semakin jarang pola yang muncul, kemudian cuaca yang diprediksi normal dan terang. Dapat dilihat pola cuaca ekstrem yang terjadi sepanjang tahun 2019 sampai 2022, cuaca ekstrem yang terjadi adalah cuaca extrem kondisi hujan. Berikut tabel dari berbagai percobaan nilai *treshlod*.

Tabel 3. Hasil percobaan berbagai treshlod

No	Nilai Treshlod	Pola Cuaca Ekstrem Yang Muncul	Prediksi Kondisi Yang Terjadi
1	28	392	Terang kondisi normal : 6, Hujan kondisi extrem : 1
2	31	211	Terang kondisi normal : 7
3	34	94	Terang kondisi normal : 7
4	37	48	Terang kondisi normal : 7
5	40	29	Terang kondisi normal : 7
6	43	21	Terang kondisi normal : 7
7	46	13	Terang kondisi normal : 7
8	49	7	Terang kondisi normal : 7

4.6. Interpretasi hasil

Hasil yang didapat penelitian ini, dari data tanggal 1 januari 2019 sampai 31 desember 2022 dengan menginputkan *treshlod value* yang diinginkan, maka dapat memberi output informasi yang sesuai dengan nilai inputan. Nilai cuaca normal yang didapat dari penelitian ini yaitu 26.84 dibulatkan menjadi 27. Semakin tinggi niai *treshlod value*, maka cuaca semakin ekstrem dan semakin jarang terjadi.

Metode yang digunakan yaitu *Long Short Term Memory*, untuk memprediksi dan membaca pola cuaca. Menentukan terjadinya cuaca ekstrem atau normal dengan cara menghitung rata – rata nilai pada setiap kolom, karena setiap atribut memiliki keterkaitan yang saling mendukung. Seperti gambar dibawah ini.

```
        Tanggal  Tavg  RH_avg  RR  ss  ff_x Kondisi_Cuaca  Rata_rata    Label
0    01-01-2019    25      89   9   1    10         HUJAN       26.8   Normal
1    02-01-2019    26      84   1   1     5       MENDUNG       23.4   Normal
2    03-01-2019    26      76   2   6     3        TERANG       22.6   Normal
3    04-01-2019    26      77   4   5     4        TERANG       23.2   Normal
4    05-01-2019    26      76   6  11     7        TERANG       25.2   Normal
```

Gambar 21. cuaca normal

Pada tanggal 01 – 01 – 2019 nilai *Tavg (temperatur rata - rata) yaitu 25 celcius, RH_avg (kelembaban) memiliki nilai 89 % menandakan sangat lembab, RR (curah hujan) memiliki nilai 9 mm, ss (intensitas sinar matahari) memiliki nilai 1 jam, dan ff_x (kecepatan angin maksimum) memiliki nilai 10 m/s.* rata – rata yang didapat adalah 26.8 maka kondisi normal. Kondisi cuaca hujan disebabkan karena nilai *RR (curah hujan) lebih tinggi dibanding dengan nilai ss (intensitas sinar matahari)* sesuai dengan sebuah jurnal dan juga berlaku sebaliknya. Didukung juga oleh nilai *RH_avg (kelembaban)* yaitu 89% sangat lembab, dan variabel lain yang berkaitan. Kemudian berikut hasil kondisi cuaca extrem.

```
Threshold Values:
          Tanggal  Rata_rata  Tavg  RH_avg   RR  ss  ff_x Kondisi_Cuaca  \
115   26-04-2019       51.4    25      88  134   6     4         Hujan
348   15-12-2019       53.4    25      87  141   8     6         Hujan
442   18-03-2020       50.0    26      91  121   8     4         Hujan
1292  16-07-2022       56.2    23      92  155   5     6         Hujan

        Label
115   Ekstrem
348   Ekstrem
442   Ekstrem
1292  Ekstrem
```

Gambar 22. cuaca extrem

Pada tanggal 26 – 04 – 2019 didapat nilai 51.4 kondisi yang terjadi adalah extrem karena diatas nilai rata – rata cuaca normal yaitu 26.84. Dengan melihat variabel *Tavg (temperatur rata - rata)* yang memiliki nilai 25 termasuk nilai normal pada sebuah temperature harian, namun dengan nilai *RR (curah hujan)* mencapai 134, curah hujan yang terjadi sangat deras dan dapat mengakibatkan banjir di wilayah yang terjadi hujan.

4.7. Keterkaitan dengan Literatur terdahulu

Penelitian ini menggunakan metode dengan akurasi terbaik menurut penelitian sebelumnya yaitu *algortima Long Short Term Memori* dengan metode *MSE* sebagai fungsi untuk meminimalkan perbedaan nilai sebenarnya dan nilai prediksi, serta mengembangkan lebih lanjut dalam pengolahan data dengan menggunakan model *RandomForestClassifier*. Dari hasil yang didapat penelitian ini mendukung dan melengkapi penelitian sebelumnya. Hasil yang didapat penelitian sebelumnya hanya menguji metode dan algoritma, pada penelitian ini dihasilkan pola cuaca dengan grafik dan hasil prediksi cuaca selama 7 hari seperti pada Tabel 4. 6 Hasil percobaan berberbagai treshlod.

4.8. Konstribusi terhadap pengetahuan

Dari Teknik yang dilakukan penelitian ini dapat memberi pengetahuan bahwa artificial intelegent mampu untuk memprediksi dan menggambarkan dengan menggunakan grafik pada bidang analisis cuaca. Penelitian ini juga menyumbangkan temuan baru berupa pola – pola yang digambarkan dengan metode algortima *Long Short Term Memory* menggunakan grafik serta menghasilkan prediksi yang cukup baik

5. KESIMPULAN DAN SARAN

Hasil yang telah didapat dari penelitian ini dan dapat dikembangkan Kembali, sebagai berikut. Kecerdasan buatan dengan algortima *Long Short Term Memory* dapat digunakan untuk Menerapkan RandomForestClassifier, pada data cuaca historis untuk memprediksi cuaca ekstrem telah menghasilkan akurasi 100%. Analisis menggunakan teknik evaluasi model MSE dengan 50 epoch dan percobaan 8 kali pada nilai threshold (28, 31, 34, 37, 40, 43, 46, 49) menunjukkan pola yang berbeda, dari rata-rata normal, yaitu 26.84 di wilayah Jawa Barat. Penelitian ini dapat dikembangkan Kembali menggunakan berbagai sumber lain dan Teknik evaluasi lain seperti *RMSE, MAE, MAPE, dan RMSE*.

DAFTAR PUSTAKA

[1] D. Manongga, U. Rahardja, I. Sembiring, N. Lutfiani, and A. B. Yadila, “Dampak Kecerdasan Buatan Bagi Pendidikan,” *ADI Bisnis Digit. Interdisiplin J.*, vol. 3, no. 2, pp. 41–55, 2022, doi: 10.34306/abdi.v3i2.792.

[2] R. L. Napitupulu and K. Handoko, “Perancangan Sistem Peramalan Cuaca Berbasis Logika Fuzzy Di Kota Batam,” *Comasie*, vol. 6, no. 5, pp. 37–44, 2022.

[3] A. M. Siregar, “Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning,” *Petir*, vol. 13, no. 2, pp. 138–147, 2020, doi: 10.33322/petir.v13i2.998.

[4] A. Prakasa and F. D. Utami, “BMKG The Integrated Weather Radar Information System of BMKG,” *J. Telecomunication, Electron. Control Eng.*, vol. 1, no. 2, pp. 86–96, 2019.

[5] J. Journal and O. F. Mathematics, “Peramalan Data Cuaca Ekstrim Indonesia Menggunakan Model ARIMA dan Recurrent Neural Network,” vol. 5, no. 1, pp. 230–242, 2023.

[6] N. M. M. Candra Devi, I. P. A. Bayupati, and N. K. A. Wirdiani, “Prediksi Curah Hujan Dasarian dengan Metode Vanilla RNN dan LSTM untuk Menentukan Awal Musim Hujan dan Kemarau,” *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 3, p. 405, 2022, doi: 10.26418/jp.v8i3.56606.

[7] Setiawan, “Keterampilan Berfikir Kritis Pada Pembelajaran IPA Menggunakan Model Pembelajaran Radec,” *Justek J. Sains dan Teknol.*, vol. 5, no. 2, p. 133, 2022.

[8] A. S. Bayangkari Karno, “Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) Dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python.,” *Ultim. InfoSys J. Ilmu Sist. Inf.*, vol. 11, no. 1, pp. 1–7, 2020, doi: 10.31937/si.v9i1.1223.

[9] M. Rizki, S. Basuki, and Y. Azhar, “Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory(LSTM) Untuk Prediksi Curah Hujan Kota Malang,” *J. Repos.*, vol. 2, no. 3, pp. 331–338, 2020, doi: 10.22219/repositor.v2i3.470.

[10] T. A. Faisal Muhammad and M. I. Irawan, “Implementasi Long Short-Term Memory (LSTM) untuk Prediksi Intensitas Curah Hujan (Studi Kasus: Kabupaten Malang),” *J. Sains dan Seni ITS*, vol. 12, no. 1, 2023, doi: 10.12962/j23373520.v12i1.106892.