



**KLASIFIKASI TIME-SERIES DATA HUJAN HARIAN DENGAN
METODE DEEP LEARNING 1D-CNN DAN LSTM PADA STASIUN
SOEKARNO HATTA, TANGERANG, INDONESIA**

TUGAS AKHIR

Rizky Widiputro

41518010039

UNIVERSITAS
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2022**



**KLASIFIKASI TIME-SERIES DATA HUJAN HARIAN DENGAN
METODE DEEP LEARNING 1D-CNN DAN LSTM PADA STASIUN
SOEKARNO HATTA, TANGERANG, INDONESIA**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Rizky Widiputro

41518010039

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA

2022

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518010039

Nama : Rizky Widiputro

Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode
Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno
Hatta, Tangerang, Indonesia

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.



UNIVERSITAS
MERCU BUANA

Jakarta, 06 Juli 2022



Rizky Widiputro

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Rizky Widiputro
NIM : 41518010039
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 06 Juli 2022



Rizky Widiputro

iii

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Rizky Widiputro
NIM : 41518010039
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis	Status
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi	Diajukan ✓
		Jurnal Nasional Terakreditasi ✓	
		Jurnal International Tidak Bereputasi	Diterima
		Jurnal International Bereputasi	
Disubmit/dipublikasikan di :	Nama Jurnal : Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)		
	ISSN : 2580-0760		
	Link Jurnal : https://jurnal.iaii.or.id		
	Link File Jurnal Jika Sudah di Publish		

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 07 Juli 2022



Rizky Widiputro

LEMBAR PERSETUJUAN PENGUJI

NIM : Rizky Widiputro
Nama : 41518010039
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



LEMBAR PERSETUJUAN PENGUJI

NIM : Rizky Widiputro
Nama : 41518010039
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



(Vina Ayumi, S.Kom., M.Kom)

LEMBAR PERSETUJUAN PENGUJI

NIM : Rizky Widiputro
Nama : 41518010039
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022



LEMBAR PENGESAHAN

NIM : 41518010039
Nama : Rizky Widiputro
Judul Tugas Akhir : Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Sockarno Hatta, Tangerang, Indonesia

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Juli 2022

Menyetujui,

UNIVERSITAS
MERCU BUANA
(Wenny Salamah, ST, MMSL)
Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)

Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)

Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur saya panjatkan kehadirat Allah SWT yang telah memberikan segala rahmat dan karunia-Nya, sehingga saya dapat menyelesaikan Tugas Akhir yang berjudul “Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia” dalam jangka waktu yang sudah ditentukan. Tanpa pertolongan-NYA, mungkin saya tidak akan sanggup menyelesaikan dengan baik. Laporan Tugas Akhir ini dibuat sebagai syarat untuk LULUS sebagai sarjana Ilmu Komputer dari Universitas Mercu Buana. Penulis menyadari dalam pembuatan Tugas Akhir ini tidak terlepas dari bantuan dan bimbingan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Keluarga saya, khususnya kedua orang tua saya yang telah menyemangati, mendukung dan selalu mendoakan yang terbaik untuk proses meraih gelar sarjana. Serta selalu percaya bahwa saya bisa menyelesaikan kuliah dengan baik.
2. Bapak Dr. Harwikarya, MT selaku Dosen Pembimbing Akademik yang telah memantau dan membantu studi saya hingga akhir.
3. Ibu Umniy Salamah, ST., MMSI, selaku dosen Pembimbing Tugas Akhir yang telah memberikan masukan saat bimbingan dan meluangkan waktu sebagian besarnya untuk melakukan bimbingan dalam penyusunan tugas akhir ini hingga selesai.
4. Seluruh Dosen Program Studi Teknik Informatika yang sudah memberikan ilmu yang bermanfaat selama kuliah berlangsung. Memberi kesempatan untuk belajar, berkarya dan juga berkembang.
5. Seluruh Staff Administrasi dan Tata Usaha yang telah banyak membantu dan memberikan kemudahan, terima kasih atas semua pelayanan dan arahannya.
6. Sahabat dan kerabat, yang telah memberikan semangat, dukungan dan doa sehingga saya bisa melewati dan menyelesaikan Tugas Akhir ini dan mendapatkan gelar sarjana.
7. Pihak-pihak dan personal yang tidak dapat disebut satu per satu yang terlibat dalam pembuatan Tugas Akhir ini sehingga dapat selesai dengan baik.

Akhir kata, penulis berharap Tugas Akhir ini dapat menjadi referensi untuk dikembangkan lebih baik lagi kedepannya, karena penulis menyadari hasil Tugas Akhir ini masih jauh dari sempurna. Masih terdapat kekurangan dalam eksperimen, cara penjelasan maupun kekeliruan penulisan. Untuk itu, kritik dan saran pembaca sangat dihargai dan diharapkan. Semoga Tugas Akhir ini dapat memberikan manfaat bagi para pembaca.

Jakarta, 21 Juni 2022

Penulis



DAFTAR ISI

HALAMAN SAMPUL.....	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... iii	
SURAT PERNYATAAN LUARAN TUGAS AKHIR.....	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT	x
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
NASKAH JURNAL	1
KERTAS KERJA.....	9
BAB 1. LITERATUR REVIEW	10
BAB 2. ANALISIS DAN PERANCANGAN.....	25
BAB 3. SOURCE CODE	29
BAB 4. DATASET.....	38
BAB 5. TAHAPAN EKSPERIMEN.....	40
BAB 6. HASIL SEMUA EKSPERIMEN.....	50
DAFTAR PUSTAKA	57

LAMPIRAN DOKUMEN HAKI..... 60

LAMPIRAN KORESPONDENSI 62



NASKAH JURNAL

Terakreditasi SINTA Peringkat 2

Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Nomor: 158/E/KPT/2021
masa berlaku mulai Volume 5 Nomor 2 Tahun 2021 sampai Volume 10 Nomor 1 Tahun 2026

Terbit online pada laman web jurnal: <http://jurnal.iaii.or.id>



JURNAL RESTI

(Rekayasa Sistem dan Teknologi Informasi)

Vol. 6 No. x (2022) x - x

ISSN Media Elektronik: 2580-0760

Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia

Daily Rainfall Time-Series Classification with Deep Learning Method 1D-CNN and LSTM on Soekarno Hatta Station, Tangerang, Indonesia

Rizky Widiputro¹, Umniy Salamah²

Teknik Informatika, Ilmu Komputer, Universitas Mercu Buana

¹41518010039@student.mercubuana.ac.id, ²umniy.salamah@mercubuana.ac.id

Abstract

Weather has very close relationship with people's life, economic, transportation, agriculture and many other aspects. Rainfall's change daily leads weather to uncertain every day. Meteorology, Climatology, and Geophysical Agency (BMKG) is non-departmental government agency. So far BMKG predicts daily rainfall based on subjective method, that the prediction made of consideration and evaluation forecaster. Therefore, methods are need to predict daily rainfall. This research used Long Short-Term Memory (LSTM) and One Dimensional Convolutional Neural Network (1D-CNN) method to classify rainfall from time-series dataset. The method predicts rainfall on day 6 with data inputs from weather's data 5 days before. To get a good model, this research used hyperparameter tuning with Gridsearch CV. The results are 1D-CNN performs well than the LSTM, although the accuracy is not much different. 1D-CNN needs time to model training less than LSTM. 1D-CNN has accuracy score 0.781 and macro f1 score 0.378, and for LSTM has accuracy score 0.758 and macro f1 score 0.353.

Keywords: Rainfall, Time-series Classification, 1D-CNN, LSTM, Deep Learning

Abstrak

Cuaca memiliki hubungan sangat erat kepada kehidupan manusia, ekonomi, transportasi, pertanian dan aspek lainnya. Perubahan curah hujan menyebabkan cuaca menjadi tak menentu setiap harinya. Badan Meteorologi, Klimatologi dan Geofisika (BMKG) merupakan lembaga Pemerintahan Non Kementrian. Selama ini BMKG memprediksi curah hujan menggunakan metode subyektif, yaitu prediksi yang dibuat berdasarkan pertimbangan atau penilaian prakirawan. Oleh karena itu dibutuhkan metode untuk memprediksi curah hujan. Penelitian ini menggunakan metode Long Short-Term Memory (LSTM) dan One Dimensional Convolutional Neural Network (1D-CNN) untuk mengklasifikasikan curah hujan pada data time-series. Pada penelitian ini memprediksi kelas curah hujan hari ke-6 dengan data cuaca interval 5 hari sebelumnya. Untuk mendapatkan model yang baik digunakan hyperparameter tuning dengan Gridsearch CV. Hasilnya 1D-CNN menghasilkan performa lebih baik dibandingkan dengan LSTM walaupun tidak jauh dari nilai akurasi. 1D-CNN membutuhkan waktu lebih cepat dibandingkan dengan LSTM. 1D-CNN memiliki akurasi 0.781 dan macro f1 score 0.378 sedangkan LSTM mempunyai nilai akurasi 0.758 dan macro f1 score 0.353.

Kata kunci: Curah Hujan, Klasifikasi Deret Waktu, 1D-CNN, LSTM, Deep Learning

Universitas Mercu Buana

<https://lib.mercubuana.ac.id/>

1. Pendahuluan

Cuaca memiliki hubungan sangat erat kepada kehidupan manusia, ekonomi, transportasi, pertanian dan aspek lainnya. Perubahan curah hujan menyebabkan cuaca menjadi tak menentu setiap harinya. Curah hujan juga menjadi penyebab terjadinya banjir, gangguan transportasi seperti bandar udara, bidang pertanian, wisata dan masih banyak lainnya.

Badan Meteorologi, Klimatologi dan Geofisika (BMKG) merupakan lembaga Pemerintahan Non Kementrian yang menyediakan informasi Informasi prakiraan cuaca, maritim, penerbangan, iklim, kualitas udara, gempabumi, tsunami dan tanda waktu di Indonesia. Selama ini BMKG memprediksi curah hujan menggunakan metode subyektif, yaitu prediksi yang dibuat berdasarkan pertimbangan atau penilaian prakirawan.

Penelitian yang dilakukan sebelumnya dalam prediksi limpahan air hujan dengan metode LSTM dan 1D CNN [1]. Hasil yang ditunjukkan memberikan model CNN sedikit lebih baik dibandingkan dengan LSTM dalam memprediksi limpahan air hujan, walaupun 2 model tersebut sama bagusnya. Begitu pula dengan prediksi cuaca (temperatur, kelembapan relatif dan kecepatan angin)[2] memberikan hasil yang baik untuk hybrid model dari 1D-CNN dan Bi-LSTM. Pada penelitian [3] dalam memprediksi stock market NKE dan GOOGL algoritma LSTM menunjukkan hasil keakuratan model yang baik untuk data time series. Pada penelitian [4] penggunaan data iklim harian dalam memprediksi curah hujan telah dibuktikan lebih memberikan keakuratan model dibandingkan dengan data El-Nino dan IOD data, penelitian ini menggunakan data time-series hanya atribut curah hujan saja. Pada penelitian [5] memprediksi kecepatan angin dari data *time-series* menunjukkan LSTM lebih baik dalam nilai RMSE dibandingkan dengan 1D-CNN.

Parameter yang digunakan adalah data cuaca harian, pada penelitian [6] parameter kelembapan, temperature dan kecepatan angin menunjukkan parameter tersebut dapat memprediksi curah hujan yang baik. Pada penelitian [7] data *time-series* menunjukkan nilai MSE yang paling kecil.

Oleh karena itu penelitian ini menggunakan 1D-CNN dan LSTM sebagai pembanding model untuk memprediksi kelas curah hujan dengan data interval *time-series*. Diharapkan penelitian ini dapat menjadi metode alternatif dalam

memprediksi curah hujan harian di Stasiun Meteorologi Soekarno Hatta. Untuk mencari hyperparameter optimal pada tiap model maka dilakukan hyperparameter tuning dengan nilai yang dievaluasi adalah macro f1 score. Untuk mengetahui performa tiap model maka akan ditampilkan classification report yang mencakup accuracy, precision, recall dan f1 score.

2. Metode Penelitian

2.1. Pengumpulan Data

Penelitian ini menggunakan data cuaca harian yang diunduh dari website data online BMKG (<https://dataonline.bmkg.go.id>). Dataset adalah data cuaca harian perbulan dari Januari 2016 sampai Desember 2020 pada Stasiun Meteorologi Soekarno Hatta, Tangerang, Indonesia. Dataset berjumlah 1,346 data dan 11 atribut yang dapat dilihat pada Tabel 1.

Tabel 1. Atribut Dataset

Atribut	Keterangan
Tanggal	Waktu Pengukuran yang berisi tanggal hari, bulan dan tahun.
Tn	Temperatur suhu minimal (°C)
Tx	Temperatur suhu maksimal (°C)
Tavg	Temperatur suhu rata-rata (°C)
RH_avg	Persentase kelembapan rata-rata (%)
RR	Nilai besarnya curah hujan dalam satuan (mm)
ss	Lamanya penyinaran matahari dalam (jam)
ff_x	Nilai kecepatan angin maksimum dalam satuan (m/s)
ddd_x	Arah angin pada kecepatan maksimum dalam derajat (°)
ff_avg	Nilai kecepatan angin rata-rata dalam satuan (m/s)
ddd_car	Arah angin yang paling sering atau terbanyak.



Gambar 1. Tampilan Data Online BMKG

Curah hujan akan dibagi menjadi 3 kelas yang dapat dilihat pada Tabel 2.

Tabel 2. Kategori Curah Hujan

Category	Rainfall
Tidak Hujan	0 mm / hour or 0 - 5 mm / day
Light Rain	1 - 5 mm / hour or 5 - 20 mm / day
Rain	>5 mm / hour or >20 mm / day

2.2. Data Preprocessing

Pada proses ini data dibersihkan dari data nul. Data null memberikan dataset menjadi tidak bisa di proses. Setelah itu menghapus atribut yang tidak diperlukan [8]. Penelitian ini sudah menentukan 7 atribut yang digunakan yaitu Temperatur minimum ($^{\circ}\text{C}$)/Tn, Temperatur maximum ($^{\circ}\text{C}$)/Tx, Temperatur rata-rata ($^{\circ}\text{C}$)/Tavg, Kelembapan rata-rata (%) /RH_avg, Lamanya penyinaran matahari (jam)/ss, Kecepatan angin rata-rata (m/s)/ff_avg dan Curah hujan (mm)/RR.

Proses dilanjutkan dengan membagi atau *split data* dengan komposisi data training 80% dan data test 20%[9]. Setelah dibagi menjadi data parameter dan data label maka data label dapat dikategorikan menjadi 3 kelas yang dapat dilihat pada Tabel 2, yang nantinya akan menjadi kelas sunny = 0, light rain = 1 dan rain = 1.

Data-data yang ada memiliki satuan yang berbeda beda, oleh karena itu data harus dinormalisasi [10]. Jenis normalisasi yang dipakai adalah *z-score normalization* dimana juga mengurangi kompleksitas data. Rumus yang dipakai dapat dilihat pada rumus 1.

$$X_{new} = \frac{X - \mu}{\sigma} \quad (1)$$

Dengan X_{new} adalah nilai setelah normalisasi, X adalah nilai sebelum dinormalisasi, μ adalah nilai rata-rata data atribut, σ adalah nilai standar deviasi data atribut. Data yang telah dinormalisasi dapat dilihat pada Gambar 2.

	0	1	2	3	4	5	6
0	-0.212790	-0.175772	0.250545	0.801472	0.632032	-0.091618	-0.465939
1	0.770725	0.764010	0.355769	0.334814	-0.368014	-0.758830	-0.465939
2	0.770725	0.294119	0.671444	0.801472	-0.368014	0.175267	-0.465939
3	0.770725	-0.254088	-0.591254	1.268130	-0.368014	-0.158339	-0.465939
4	0.770725	0.920540	0.987118	0.334814	-0.368014	0.675676	0.922843
...
1071	-0.606196	-3.465010	-2.485301	1.734788	2.618138	-1.826369	-0.465939
1072	-0.311142	-0.489033	-1.117378	0.957025	0.387266	-1.826369	0.922843
1073	0.180616	0.059173	0.460994	0.023709	-0.123248	-0.024896	0.922843
1074	1.262483	-1.585445	-0.801704	1.423683	-0.298081	1.109364	-0.465939
1075	0.475670	-0.332403	0.460994	0.179261	0.072566	-1.192517	0.922843

Gambar 2. Dataset Setelah Normalisasi

Data label setelah dikategorikan menjadi 3 kelas diubah menjadi *one-hot vector* untuk dapat terlihat *probability distribution* yang dievaluasi dengan loss. Hasil *one-hot vector* dapat dilihat pada Tabel 3.

Tabel 3. One-hot Encoding Class

Class	One-hot Vector
0	[1, 0, 0]
1	[0, 1, 0]
2	[0, 0, 1]

Setelah itu data dilakukan *windowing* atau membuat data interval/time step dengan 5 hari data cuaca sebagai input dan data curah hujan hari ke-6 sebagai data label.

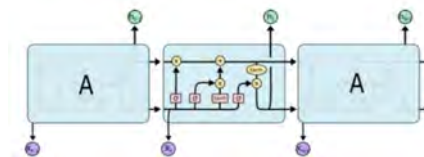
Tabel 4. Dataset setelah windowing

Variable Input	Variable Label
D_1, D_2, D_3, D_4, D_5	CH_6
D_2, D_3, D_4, D_5, D_6	CH_7
.....
D_{n-5}, \dots, D_{n-1}	CH_n

Dimana D_1 adalah data cuaca hari ke-1, dan CH_6 adalah data curah hujan hari ke-6.

2.3. Long Short-Term Memory (LSTM)

Recurrent Neural Network (RNN) adalah sebagian dari deep learning yang mempunyai koneksi yang berulang pada tiap *cell* didalam *hidden layer* [11]. LSTM merupakan modifikasi dari metode RNN [12]. LSTM bertujuan untuk mengurangi *long-term dependency* dari metode RNN dengan menggunakan *memory cells* untuk mempertahankan *state* dalam waktu lama[13]. Arsitektur dari algoritma LSTM dapat dilihat pada Gambar 3[14].



Gambar 3. Arsitektur LSTM

Untuk memahami lebih dalam mengenai LSTM, langkah-langkahnya harus dijelaskan. Berdasarkan [4] langkah-langkah untuk pembentukan model LSTM ditunjukkan pada rumus 3 – rumus 8.

$$\sigma(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

Dimana f_t adalah forget gate, σ adalah fungsi sigmoid, x_t adalah input vektor, W_f adalah weight dari forget gate, h_{t-1} adalah nilai output dari model sebelumnya. Selanjutnya model memilih informasi apa yang akan disimpan. Proses selanjutnya adalah pada input gate.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

$$\bar{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \bar{C}_t \quad (6)$$

Dimana i_t adalah nilai input gate, \bar{C}_t adalah nilai cell state baru yang ditambahkan, C_t adalah nilai cell state. Dan yang terakhir adalah output gate.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \cdot \tanh C_t \quad (8)$$

Dimana o_t adalah nilai output gate, h_t adalah nilai output nya. Setelah dilakukan hyperparameter tuning dengan Gridsearch CV maka Model pada penelitian ini dapat dilihat pada Gambar 4.

Layer (type)	Output Shape	Param #
lstm_4 (LSTM)	(None, 5, 16)	1536
dropout_10 (Dropout)	(None, 5, 16)	0
lstm_5 (LSTM)	(None, 128)	74240
dropout_11 (Dropout)	(None, 128)	0
dense_12 (Dense)	(None, 64)	8256
dropout_12 (Dropout)	(None, 64)	0
dense_13 (Dense)	(None, 3)	195

 Total params: 84,227
 Trainable params: 84,227
 Non-trainable params: 0

Gambar 4. Model LSTM

Parameter yang digunakan untuk dropout 1 dan 2 dengan rate 0.2 dan dropout 3 dengan rate 0.5. Optimizer model menggunakan adam dengan learning rate 0.01 dan digunakan exponential decay. Untuk batch size digunakan 32 data per epochnya. Setelah itu digunakan earllystopping untuk memberhentikan training jika validasi loss tidak turun dalam 5 epoch. Nilai loss yang digunakan adalah crossentropy loss atau negative log loss.

2.4. One-Dimensional Convolutional Neural Network (1D-CNN)

1D-CNN adalah metode yang termasuk kedalam deep learning[15]. CNN adalah metode yang dikembangkan berdasarkan ide dari *local connectivity* [1]. 1D-CNN menunjukkan kinerja yang baik dalam melakukan analisis data time series.

CNN terdiri dari beberapa lapisan dimana strukturnya dirancang berdasarkan kasus. Lapisan pertama yang terlibat dalam CNN adalah lapisan konvolusi yang terdiri dari neuron dalam bentuk matriks filter. Pergeseran matriks filter akan direpresentasikan sebagai produk titik atau proses konvolusi antara input dan kernel dengan ukuran tertentu. Kemudian, proses akan menghasilkan output atau disebut peta aktivasi atau peta fitur. Lapisan (layer) adalah lapisan Rectified Linear Unit (ReLU) yang mengubah nilai negatif menjadi nol dan mempertahankan nilai positif. Layer ketiga adalah Pooling yang terdiri dari filter dengan ukuran dan stride tertentu. Pada pooling layer, operasi yang biasanya digunakan adalah operasi maksimum (max pooling) dan rata-rata (average pooling). Tujuan penggunaan pooling layer adalah untuk memperkecil dimensi dari peta fitur dan akibatnya meningkatkan kecepatan komputasi. Proses konvolusi [5] dari filter ke- i dapat ditulis pada rumus 9.

$$h_{ij}^{(k)} = \emptyset \left(\sum_{p=1}^{C^{(k-1)}} \sum_{q=1}^H W_{ipq}^{(k)} x_{p,j+q}^{(k)} + b_i^{(k)} \right) \quad (9)$$

Dimana $W_{ipq}^{(k)}$ adalah weight dari filter, $b_i^{(k)}$ adalah bias dari filter, H adalah nomor dari filter size, x adalah input dan \emptyset adalah fungsi aktivasi. Setelah dilakukan hyperparameter tuning dengan Gridsearch CV maka Model pada penelitian ini dapat dilihat pada Gambar 5.

Layer (type)	Output Shape	Param #
conv1d_10 (Conv1D)	(None, 5, 8)	120
batch_normalization_10 (Batch Normalization)	(None, 5, 8)	32
conv1d_11 (Conv1D)	(None, 4, 32)	544
max_pooling1d_5 (MaxPooling1D)	(None, 2, 32)	0
batch_normalization_11 (Batch Normalization)	(None, 2, 32)	128
flatten_5 (Flatten)	(None, 64)	0
dense_16 (Dense)	(None, 64)	4160
dropout_14 (Dropout)	(None, 64)	0
dense_17 (Dense)	(None, 3)	195

 Total params: 5,179
 Trainable params: 5,099
 Non-trainable params: 80

Gambar 5. Model 1D-CNN

Penggunaan Batch Normalization bertujuan untuk mempercepat proses optimasi parameter[16]. Parameter yang digunakan untuk dropout dengan rate 0.5. Optimizer model menggunakan adam dengan learning rate 0.01 dan digunakan exponential decay. Untuk batch size digunakan 32 data per epochnya. Setelah itu digunakan earlystopping untuk memberhentikan training jika validasi loss tidak turun dalam 5 epoch. Nilai loss yang digunakan adalah crossentropy loss atau negative log loss.

2.5. Cross-Entropy Loss

Negative log loss atau cross-entropy loss adalah *loss function* yang biasa digunakan untuk training *multi-class classification*[17]. Berikut adalah persamaan yang digunakan.

$$L_{CE} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^l y_{ij} \ln \left(\frac{e^{s_{y_j}}}{\sum_{k=1}^m e^{s_k}} \right) \quad (10)$$

Dimana n adalah jumlah sampel, l jumlah elemen pada dimensi label one-hot vector, y_{ij} adalah nilai label pada one-hot vektor elemen ke- j pada sampel ke- i , m adalah Jumlah elemen pada dimensi nilai prediksi output layer, s_{y_j} adalah Nilai prediksi pada dimensi output layer elemen ke- j , s_k adalah Nilai prediksi pada dimensi output layer elemen ke- k .

2.6. Classification report

Classification report adalah sebuah laporan mengenai hasil model. Hasil model pada prediksi data mencakup nilai *accuracy*, *precision*, *recall* dan *f1 score* tiap kelas label.

$$precision = \frac{TruePositive}{TruePositive+FalsePositi} \quad (11)$$

$$Recall = \frac{TruePositive}{TruePositive+FalseNegati} \quad (12)$$

$$accuracy = micro\ f1\ score = \frac{TP}{TP+FP+T+FN} \quad (13)$$

$$f1sm = \frac{1}{n} \sum_{i=1}^n 2 \cdot \frac{precision_i \cdot recall_i}{precision_i + recall_i} \quad (14)$$

$$f1sw = \frac{1}{n} \cdot \frac{1}{S_t} \cdot \sum_{i=1}^n 2 S_i \cdot \frac{precision_i \cdot recall_i}{precision_i + recall_i} \quad (15)$$

Dimana $f1sm$ adalah macro f1 score, $f1sw$ adalah weighted f1 score, n adalah jumlah sampel, S_t adalah total sampel semua kelas dan S_i adalah total sampel kelas i .

3. Hasil dan Pembahasan

3.1. One-Dimensional Neural Network (1D-CNN)

Eksperimen yang pertama dilakukan adalah penggunaan hyperparameter tuning dengan Gridsearch CV. Hyperparameter tuning Gridsearchcv memakai cross-validation(cv) 5. Dan untuk mengevaluasi performa dari cv dari data yg dimasukkan penelitian ini menggunakan nilai macro f1 score. Untuk perancangan model menjadi dasar penggunaan gridsearch cv terdapat pada Gambar 5. Setelah menggunakan hyperparameter tuning didapatkan nilai macro f1 score terbaik yang dicoba dan dengan parameter nya.

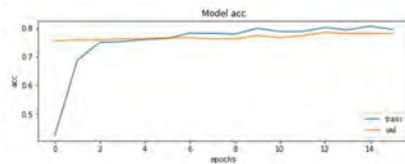
Setelah didapatkan parameter terbaik maka dimasukkan kedalam model 1D-CNN. Selanjutnya untuk mendapatkan akurasi yg lebih baik maka digunakan exponential decay pada optimizer adam dengan learning rate 0.01, decay steps 136 dan decay rate 0.96.

Lalu untuk menghindari model overfitting terhadap data digunakan earlystopping. Earlystopping adalah suatu fungsi pada keras tensorflow yang digunakan untuk memberhentikan epoch atau training model jika suatu nilai tidak berkurang atau bertambah. Pada penelitian ini digunakan earlystopping dengan memonitor nilai loss pada data validasi, jika nilai loss tidak berkurang setelah 5 epoch maka training model dihentikan.

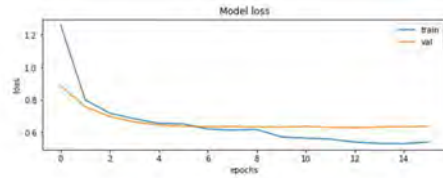
```
Epoch 10/100: 11.19ms/step - loss: 0.2094 - acc: 0.4212 - val_loss: 0.8884 - val_acc: 0.7547
Epoch 11/100: 11.19ms/step - loss: 0.2097 - acc: 0.4212 - val_loss: 0.7968 - val_acc: 0.7547
Epoch 12/100: 11.19ms/step - loss: 0.2108 - acc: 0.4205 - val_loss: 0.8083 - val_acc: 0.7547
Epoch 13/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8463 - val_acc: 0.7423
Epoch 14/100: 11.19ms/step - loss: 0.2093 - acc: 0.4200 - val_loss: 0.8454 - val_acc: 0.7423
Epoch 15/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 16/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 17/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 18/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 19/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 20/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 21/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 22/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 23/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 24/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 25/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 26/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 27/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 28/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 29/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
Epoch 30/100: 11.19ms/step - loss: 0.2100 - acc: 0.4200 - val_loss: 0.8377 - val_acc: 0.7500
```

Gambar 6. Process Training 1D-CNN

Dapat dilihat jika waktu yang diperlukan untuk training model 1D-CNN berkisar di waktu 4-5 ms per step, dan jika dilihat pada Gambar 5 total params adalah 5,179.



Gambar 7. Visualisasi Accuracy Training 1D-CNN



Gambar 8. Visualisasi Loss Training 1D-CNN

Selanjutnya dilakukan prediksi terhadap data test yang sebelumnya telah dipersiapkan dari awal dataset dibagi. Hasil *classification report* dapat dilihat pada Gambar 9.

Classification report CNN :

	precision	recall	f1-score	support
tidak hujan	0.792	0.995	0.882	203
light Rain	0.571	0.093	0.160	43
Rain	0.333	0.053	0.091	19
accuracy			0.781	265
macro avg	0.566	0.380	0.378	265
weighted avg	0.723	0.781	0.708	265

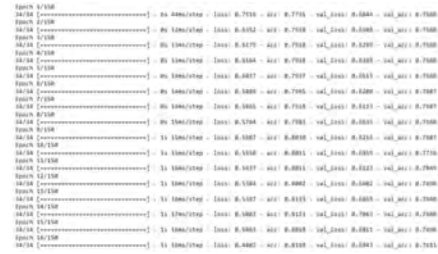
Gambar 9. Classification Report 1D-CNN

Hasil klasifikasi dapat dilihat dari precision tiap kelas berturut-turut 0.792, 0.571 dan 0.333, nilai accuracy sebesar 0.781 dan macro f1 score 0.378. Pada model 1D-CNN dapat mengklasifikasi 3 kelas walaupun nilai precision pada kelas rain hanya bernilai 0.33.

3.2. Long Short-Term Memory (LSTM)

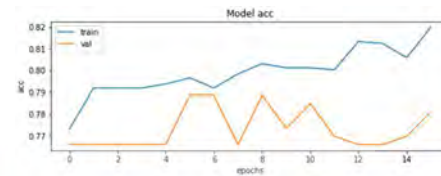
Eksperimen pada model LSTM dimulai dari hyperparameter tuning menggunakan gridsearch cv. Gridsearch cv yang digunakan juga sama dengan 1D-CNN yaitu pada evaluasi performa menggunakan macro f1 score. Maka hasil macro f1 score dengan parameter terbaik didapatkan.

Setelah didapatkan parameter terbaik maka dimasukkan kedalam model LSTM. Selanjutnya untuk mendapatkan akurasi yg lebih baik maka digunakan exponential decay pada optimizer adam dengan learning rate 0.01, decay steps 136 dan decay rate 0.96. Untuk menghindari model overfitting terhadap data digunakan earllystopping. Earllystopping adalah suatu fungsi pada keras tensorflow yang digunakan untuk memberhentikan epoch atau training model jika suatu nilai tidak berkurang atau bertambah. Pada penelitian ini digunakan earllystopping dengan memonitor nilai loss pada data validasi, jika nilai loss tidak berkurang setelah 5 epoch maka training model dihentikan

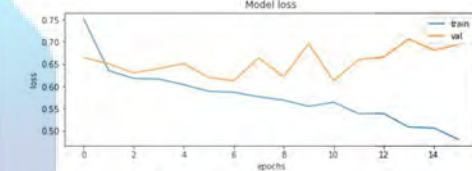


Gambar 10. Proses Training LSTM

Dapat dilihat jika waktu yang diperlukan untuk training model LSTM berkisar di waktu 12 ms – 1s 15ms per step, dan jika dilihat pada Gambar4 total params adalah 84,227.



Gambar 11. Visualisasi Accuracy Training LSTM



Gambar 12. Visualisasi Loss Training LSTM

Terdapat perbedaan yang terlihat pada visualisasi accuracy dan loss kedua model. Selanjutnya model memprediksi data test dan menghasilkan *classification report* yang ada pada.

Classification report LSTM :

	precision	recall	f1-score	support
tidak hujan	0.799	0.961	0.872	203
light Rain	0.286	0.140	0.187	43
Rain	0.000	0.000	0.000	19
accuracy			0.758	265
macro avg	0.362	0.367	0.353	265
weighted avg	0.659	0.758	0.699	265

Gambar 13. Classification Report LSTM

Hasil klasifikasi dapat dilihat dari precision tiap kelas berturut-turut 0.799, 0.286 dan 0, nilai accuracy sebesar 0.758 dan macro f1 score 0.353. Pada model 1D-CNN dapat mengklasifikasi 3 kelas pada LSTM hanya mengklasifikasi 2 kelas saja, sedangkan untuk kelas rain LSTM tidak bisa mengklasifikasikan data.

4. Kesimpulan

Berdasarkan hasil analisis dan pengujian time-series classification pada model LSTM dan 1D-CNN didapatkan parameter yang paling baik

dari proses hyperparameter tuning menggunakan Gridsearch CV. Perbandingan model 1D-CNN dan LSTM yang pertama adalah waktu training, 1D-CNN memerlukan waktu lebih baik daripada LSTM dengan waktu 4-5 ms per step sedangkan LSTM memerlukan 12ms – 1s 15ms per step. 1D-CNN juga unggul dalam nilai akurasi dan f1 score, 1D-CNN memiliki akurasi 0.781 dan macro f1 score 0.378 sedangkan LSTM mempunyai nilai akurasi 0.758 dan macro f1 score 0.353. Pada model LSTM belum bisa memprediksi kelas rain dengan nilai precision tiap kelas 0.79, 0.28, 0 sedangkan model 1D-CNN punya nilai precision tiap kelas 0.79, 0.57, 0.33. Nilai precision yang kecil pada kelas light rain dan rain salah satunya disebabkan karena imbalance dataset atau dataset dengan kelas yang tidakimbang, hal ini dikarenakan pada daerah Stasiun Meteorologi Soekarno Hatta lebih banyak hari tidak hujan dibandingkan dengan hujan. Dapat disimpulkan dari penelitian ini model CNN dapat dijadikan sebagai alternatif dalam memprediksi intensitas curah hujan pada Stasiun Meteorologi Soekarno Hatta. Untuk penelitian selanjutnya dapat digunakan beberapa time step berbeda agar diuji bagaimana performa model dengan dataset yang berbeda time stepnya. Dan penambahan atribut atau variabel input seperti data dari stasiun meteorologi berbeda dan arah angin.

Daftar Rujukan

- [1] S. P. Van, H. M. Le, D. V. Thanh, T. D. Dang, H. H. Loc, and D. T. Anh, "Deep learning convolutional neural network in rainfall-runoff modelling," *J. Hydroinformatics*, vol. 22, no. 3, pp. 541–561, 2020, doi: 10.2166/hydro.2020.095.
- [2] Q. Fu, D. Niu, Z. Zang, J. Huang, and L. Diao, "Multi-stations' weather prediction based on hybrid model using 1D CNN and Bi-LSTM," *Chinese Control Conf. CCC*, vol. 2019-July, pp. 3771–3775, 2019, doi: 10.23919/ChiCC.2019.8866496.
- [3] A. Moghar and M. Hamiche, "Stock Market Prediction Using LSTM Recurrent Neural Network," in *Procedia Computer Science*, 2020, vol. 170, pp. 1168–1173, doi: 10.1016/j.procs.2020.03.049.
- [4] D. Z. Haq *et al.*, "Long Short-Term Memory Algorithm for Rainfall Prediction Based on El-Nino and IOD Data," in *Procedia Computer Science*, 2021, vol. 179, no. 2019, pp. 829–837, doi: 10.1016/j.procs.2021.01.071.
- [5] R. Fukuoka, H. Suzuki, T. Kitajima, A. Kuwahara, and T. Yasuno, "Wind Speed Prediction Model Using LSTM and 1D-CNN," *J. Signal Process.*, vol. 22, no. 4, pp. 207–210, Jul. 2018, doi: 10.2299/jsp.22.207.
- [6] D. C. R. Novitasari *et al.*, "Weather Parameters Forecasting as Variables for Rainfall Prediction using Adaptive Neuro Fuzzy Inference System (ANFIS) and Support Vector Regression (SVR)," in *Journal of Physics: Conference Series*, 2020, vol. 1501, no. 1, doi: 10.1088/1742-6596/1501/1/012012.
- [7] U. Kaimkhani, B. Naz, and S. Narejo, "Rainfall Prediction Using Time Series Nonlinear Autoregressive Neural Network," *Int. J. Comput. Sci. Eng.*, vol. 8, no. 1, pp. 30–38, 2021, doi: 10.14445/23488387/ijcse-v8i1p106.
- [8] B. T. Pham *et al.*, "Development of advanced artificial intelligence models for daily rainfall prediction," *Atmos. Res.*, vol. 237, no. January, p. 104845, 2020, doi: 10.1016/j.atmosres.2020.104845.
- [9] A. Cahyaningsih, N. P. Putra, A. Pradika, and E. Pratama, "Model Prediksi Jumlah Kumulatif Kasus COVID-19 di Indonesia Menggunakan Metode Neural Network," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 3, no. 1, pp. 76–83, 2020, doi: 10.20895/INISTA.V2I2.
- [10] Shakib Badarpura, Abhishek Jain, Aniket Gupta, and Deepali Patil, "Rainfall Prediction using Linear approach Neural Networks and Crop Recommendation based on Decision Tree," *Int. J. Eng. Res.*, vol. V9, no. 04, pp. 394–399, 2020, doi: 10.17577/ijertv9is040314.
- [11] I. R. Widiyari, L. E. Nugoho, Widyawan, and R. Efendi, "Context-based Hydrology Time Series Data for A Flood Prediction Model Using LSTM," in *Proceedings - 2018 5th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2018*, 2018, pp. 385–390, doi: 10.1109/ICITACEE.2018.8576900.
- [12] İ. Kırbas, A. Sözen, A. D. Tuncer, and F. Ş. Kazancıoğlu, "Comparative analysis and forecasting of COVID-19 cases in various European countries with ARIMA, NARNN and LSTM approaches," *Chaos, Solitons and Fractals*, vol. 138, 2020, doi: 10.1016/j.chaos.2020.110015.
- [13] E. S. Putri and M. Sadikin, "Prediksi Penjualan Produk Untuk Mengestimasi Kebutuhan Bahan Baku Menggunakan Perbandingan Algoritma LSTM dan ARIMA," *Format J. Ilm. Tek. Inform.*, vol. 10, no. 2, p. 162, 2021, doi: 10.22441/format.2021.v10.i2.007.
- [14] M. Q. Andiyantama, I. Zahira, and A. Irawan, "Prediksi Energi Listrik Kincir Angin Berdasarkan Data Kecepatan Angin Menggunakan LSTM," *JITCE (Journal Inf. Technol. Comput. Eng.)*, vol. 5, no. 01, pp. 1–7, 2021, doi: 10.25077/jitce.5.01.1-7.2021.
- [15] C. N. Ihsan, "Klasifikasi Data Radar Menggunakan Algoritma Convolutional Neural Network (CNN)," *J. Comput. Inf. Technol.*, vol. 4, no. 2, pp. 115–121, 2021, doi: http://doi.org/10.25273/doubleclick.v4i2.8188.
- [16] Rushendra, K. Ramli, N. Hayati, E. Ihsanto, T. S. Gunawan, and A. H. Halbouni, "Development of Intrusion Detection System using Residual Feedforward Neural Network Algorithm," *2021 4th Int. Semin. Res. Inf. Technol. Intell. Syst. ISRITI 2021*, pp. 539–543, 2021, doi: 10.1109/ISRITI54043.2021.9702773.
- [17] J. Xie, K. Hu, Y. Guo, Q. Zhu, and J. Yu, "On loss

functions and CNNs for improved bioacoustic signal classification,” *Ecol. Inform.*, vol. 64, no. May, p. 101331, 2021, doi: 10.1016/j.ecoinf.2021.101331.



KERTAS KERJA

Ringkasan

Kertas kerja ini berisi tentang kelengkapan material dari artikel jurnal dengan judul “Klasifikasi Time-Series Data Hujan Harian dengan Metode Deep Learning 1D-CNN dan LSTM pada Stasiun Soekarno Hatta, Tangerang, Indonesia”. Seluruh hasil penelitian Tugas Akhir yang tidak dimasukkan kedalam artikel jurnal. Pada kertas kerja ini disajikan: *literatur review*, analisis perancangan, *source code*, dataset yang digunakan, tahapan eksperimen dan seluruh hasil eksperimen.

- Bagian 1: *Literature Review* menjabarkan mengenai beberapa jurnal yang terkait dengan penelitian.
- Bagian 2: Analisis dan Perancangan menjabarkan analisis masalah dan analisis model.
- Bagian 3: *Source Code* menjabarkan kumpulan kode di setiap prosesnya mulai dari Membaca data, Preprocessing, Implementasi model, sampai pengujian dan evaluasi.
- Bagian 4: *Dataset* melampirkan pengambilan dataset.
- Bagian 5: Tahapan eksperimen berisi pengumpulan data, *Preprocessing*, Implementasi model, dan evaluasi model.
- Bagian 6: Hasil implementasi eksperimen secara keseluruhan yang mencakup jawaban dari skenario percobaan pengujian yang sudah disebutkan pada bagian 5. Bagian ini berguna untuk menemukan model prediksi terbaik yang dilihat dari beberapa pengujian berdasarkan nilai *categorical crossentropy*. Kemudian, hasil yang didapat digunakan sebagai pemodelan visualisasi kelas curah hujan untuk bahan evaluasi atas kebijakan yang diterapkan.