

**PREDIKSI SATU DIMENSI SUHU DANAU BATUR DENGAN
MENGUNAKAN PENGGABUNGAN METODE DEEP LEARNING DAN
PHYSICS INFORMED NEURAL NETWORK**

Skripsi

Oleh

**NAUFAL HAFIZD PRAYOGA
NPM 2217031003**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

ABSTRACT

PREDICTING 1-DIMENSIONAL LAKE BATUR TEMPRATURE USING DEEP LEARNING AND PHYSICS INFORMED NEURAL NETWORK COMBINED APPROACH

By

Naufal Hafizd Prayoga

Lake Batur in Bali, Indonesia, experiences periodic mass fish mortality events caused by lake turnover, where rapid mixing brings sulfur-rich bottom water to the surface. This research develops a predictive model for Lake Batur's one-dimensional water temperature profile using a hybrid approach combining Deep Learning and Physics-Informed Neural Networks (PINN). Five models were evaluated: Random Forest (RF), Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), PINN, and PINN-LSTM. The PINN-LSTM model, which integrates LSTM's temporal pattern recognition with physics-based constraints from the heat diffusion equation, achieved the best overall performance with an RMSE of 0.203°C, NSE of 0.819, and physical inconsistency of only 3.42%. This research confirms a trade-off between data-driven accuracy and physical consistency, where pure data-driven models (LSTM, CNN) achieve high accuracy but exhibit 12-14% physical inconsistency, while the pure PINN model maintains physical consistency but performs poorly (negative NSE). The optimal configuration uses 72 hours of historical data ($n_p = 72$) as input and predicts 12 hours ahead ($n_f = 12$), providing sufficient lead time for early warning systems. Only the PINN-LSTM model was capable of reproducing the critical transition from stratification to mixing events. These findings demonstrate that hybrid physics-informed deep learning approaches offer an effective solution for lake temperature prediction and early warning system development.

Keywords: PINN-LSTM, Lake Temperature Prediction, Physics-Informed Neural Network, Deep Learning, Early Warning System

ABSTRAK

PREDIKSI SATU DIMENSI SUHU DANAU BATUR DENGAN MENGUNAKAN PENGGABUNGAN METODE DEEP LEARNING DAN PHYSICS INFORMED NEURAL NETWORK

Oleh

Naufal Hafidz Prayoga

Danau Batur di Bali, Indonesia, mengalami peristiwa kematian ikan massal secara periodik yang disebabkan oleh fenomena pencampuran danau, di mana pencampuran cepat membawa air dasar yang kaya belerang ke permukaan. Penelitian ini mengembangkan model prediksi profil suhu air satu dimensi Danau Batur menggunakan pendekatan *hybrid* yang menggabungkan *Deep Learning* dan *Physics-Informed Neural Networks* (PINN). Lima model dievaluasi: *Random Forest* (RF), *Long Short-Term Memory* (LSTM), *Convolutional Neural Network* (CNN), PINN, dan PINN-LSTM. Model PINN-LSTM, yang mengintegrasikan kemampuan pengenalan pola temporal LSTM dengan batasan berbasis fisika dari persamaan difusi panas, mencapai performa terbaik secara keseluruhan dengan RMSE sebesar 0.203°C , NSE sebesar 0.819, dan ketidakkonsistenan fisika hanya 3.42%. Penelitian ini mengkonfirmasi adanya *trade-off* antara akurasi berbasis data dan konsistensi fisika, di mana model berbasis data murni (LSTM, CNN) mencapai akurasi tinggi tetapi menunjukkan ketidakkonsistenan fisika 12-14%, sedangkan model PINN murni mempertahankan konsistensi fisika tetapi memiliki performa buruk (NSE negatif). Konfigurasi optimal menggunakan 72 jam data historis ($n_p = 72$) sebagai *input* dan memprediksi 12 jam ke depan ($n_f = 12$), memberikan waktu tunggu yang cukup untuk sistem peringatan dini. Hanya model PINN-LSTM yang mampu mereproduksi transisi kritis dari stratifikasi ke peristiwa pencampuran. Temuan ini menunjukkan bahwa pendekatan *hybrid physics-informed deep learning* menawarkan solusi yang efektif untuk prediksi suhu danau dan pengembangan sistem peringatan dini.

Kata-kata kunci: PINN-LSTM, Prediksi Suhu Danau, *Physics-Informed Neural Network*, *Deep Learning*, Sistem Peringatan Dini

**PREDIKSI SATU DIMENSI SUHU DANAU BATUR DENGAN
MENGUNAKAN PENGABUNGAN METODE DEEP LEARNING DAN
PHYSICS INFORMED NEURAL NETWORK**

Oleh

NAUFAL HAFIZD PRAYOGA

Skripsi

**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

Judul Skripsi : **PREDIKSI SATU DIMENSI SUHU DANAU
BATUR DENGAN MENGGUNAKAN
PENGGABUNGAN METODE DEEP
LEARNING DAN PHYSICS INFORMED
NEURAL NETWORK**

Nama Mahasiswa : **Naufal Hafizd Prayoga**

Nomor Pokok Mahasiswa : **2217031003**

Program Studi : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**



1. Komisi Pembimbing

Dr. Aang Nuryaman S.Si., M.Si.

NIP. 197403162005011001



Furqon H. Muttaqien, S.Komp., M.Kom.

NIP. 198808092010121003



2. Ketua Jurusan Matematika

Dr. Aang Nuryaman S.Si., M.Si.

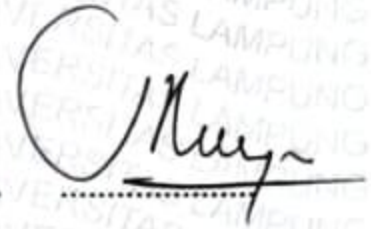
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

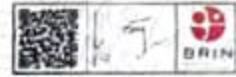
Ketua

: **Dr. Aang Nuryaman S.Si., M.Si.**



Sekretaris

: **Furqon H. Muttaqien, S.Komp.,
M.Kom.**



Penguji

Bukan Pembimbing : **Dr. Dian Kurniasari, S.Si., M.Sc.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.

NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: **28 Januari 2026**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : **Naufal Hafizd Prayoga**
Nomor Pokok Mahasiswa : **2217031003**
Jurusan : **Matematika**
Judul Skripsi : **Prediksi Satu Dimensi Suhu Danau Batur
dengan Menggunakan Penggabungan Metode
Deep Learning dan Physics Informed Neural
Network**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 28 Januari 2026

Penulis,



Naufal Hafizd Prayoga

RIWAYAT HIDUP

Penulis memiliki nama lengkap Naufal Hafizd Prayoga yang lahir di Bekasi pada tanggal 18 Mei 2004. Perjalanan pendidikan penulis dimulai dari tingkat TK hingga SMA di Bekasi. Beliau menempuh pendidikan di TK Al-Ikhlas dari Juli 2008 hingga Juni 2010, SD Muhammadiyah 47 dari Juli 2010 hingga Juni 2016, SMP Negeri 3 Bekasi dari Juli 2016 hingga Mei 2019, dan SMA Negeri 18 Bekasi dari Juli 2019 hingga April 2022. Sejak Agustus 2022, Penulis melanjutkan studi di Universitas Lampung, Rajabasa, Kota Bandar Lampung, di Fakultas Matematika dan Ilmu Pengetahuan Alam, dengan jurusan Matematika. Penulis sebagai mahasiswa aktif yang ambisius menjadi statistikawan. Penulis memiliki pengalaman dalam statistika dan analisis data, serta dikenal suka belajar hal baru, disiplin, dan mudah beradaptasi dalam lingkungan tim. Sehingga menunjukkan komitmennya terhadap pengembangan diri dan karier di bidang yang diminatinya.

Di Universitas Lampung, penulis aktif dalam kegiatan organisasi, khususnya di UKM Catur. Penulis menjabat sebagai Wakil Ketua Departemen Hubungan Masyarakat dari Agustus 2022 hingga Juni 2023, serta Ketua Seksi Acara untuk Turnamen Catur yang memperebutkan Piala UKM Catur Universitas Lampung pada November 2022. Pengalaman ini menunjukkan kemampuannya dalam manajemen dan koordinasi acara. Selain itu, penulis juga aktif dalam mengikuti beberapa kegiatan yang diselenggarakan oleh Himatika. Kegiatan tersebut adalah kegiatan MAFIA (Mentoring Akademik of Himatika), Penulis meraih penghargaan berupa mentor terbaik yang dilaksanakan pada bulan April sampai Juni 2024.

KATA INSPIRASI

“Live Life The Fullest, Make a Difference Along The Way”

(Anonymous)

“選んだ道を正解にする”

Eranda michi wo seikai ni suru

(Make the path you chose the correct answer.)

(Brigitte Bardot)

“無理せず、怠けず”

Muri sezu, namakezu

(Don't force yourself, but don't slack off.)

(Mitsuo Aida)

PERSEMBAHAN

Segala puji dan syukur saya panjatkan ke hadirat Allah SWT atas segala nikmat dan petunjuk-Nya, sehingga penyusunan skripsi ini dapat diselesaikan dengan baik dan tepat waktu. Dengan penuh rasa syukur dan kebahagiaan, saya sampaikan ucapan terima kasih saya kepada:

Ayah dan Ibuku Tercinta

Terima kasih yang sebesar-besarnya kepada kedua orang tuaku atas segala bentuk pengorbanan, doa, motivasi, restu, dan dukungan yang tiada henti. Terima kasih telah mengajarkan arti penting dari perjalanan hidup, yang menjadi bekal berharga bagi saya untuk tumbuh menjadi pribadi yang bermanfaat bagi banyak orang di masa depan.

Dosen Pembimbing dan Pembahas

Terima kasih kepada dosen pembimbing dan dosen pembahas atas segala bantuan, motivasi, arahan, serta ilmu yang begitu berharga yang telah diberikan sepanjang proses penyusunan skripsi ini.

Sahabat-sahabatku

Terima kasih kepada semua orang baik yang telah berbagi pengalaman, semangat, motivasi, serta doa, dan yang selalu memberikan dukungan dalam berbagai hal.

Almamater Tercinta

Universitas Lampung

SANWACANA

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Prediksi Satu Dimensi Suhu Danau Batur dengan Menggunakan Penggabungan Metode Deep Learning dan Physics Informed Neural Network” dengan baik, lancar, dan tepat waktu. Shalawat serta salam semoga selalu tercurah kepada Nabi Muhammad SAW.

Dalam proses penyusunan skripsi ini, penulis menerima banyak bantuan berupa bimbingan, dukungan, arahan, motivasi, dan saran dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Pak Dr. Aang Nuryaman S.Si., M.Si.. selaku Pembimbing I dan Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang telah banyak meluangkan waktunya untuk memberikan arahan, bimbingan, motivasi, saran serta dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
2. Pak Furqon Hensan Muttaqien, S.Komp., M.Kom. selaku Pembimbing II yang telah memberikan arahan, bimbingan dan dukungan kepada penulis sehingga dapat menyelesaikan skripsi ini.
3. Ibu Dr. Dian Kurniasari, S.Si., M.Sc. selaku Penguji yang telah bersedia memberikan kritik dan saran serta evaluasi kepada penulis sehingga dapat menjadi lebih baik lagi.
4. Ayah, Ibu, dan Adik Perempuan yang senantiasa memberikan do'a dan motivasi kepada penulis.

5. Ibu Prof. Dra. Wamiliana, MA., Ph.D. selaku dosen pembimbing akademik.
6. Pak Arianto Budi Santoso, Ph.D. yang telah memberikan saran berharga dalam penelitian ini serta menyediakan akses terhadap dataset Danau Batur.
7. Seluruh dosen, staff dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
8. Sahabat penulis sejak SMA diantaranya Muhammad Aqil Muzhaffar, Samuel Anju Lumbantoruan, Muhammad Iqbal Abdillah, Muhammad Hilmi Irsyad, dan Hapis yang telah banyak memberikan tawa dan selalu meluangkan waktu untuk saling membantu.
9. Pimpinan HIMATIKA Periode 2025, terimakasih banyak karena sudah menjadi pimpinan yang sangat keren luar biasa.
10. Rekan seperjuangan skripsi, diantaranya Benaya, Agustino Simatupang, Yuniar, Aisyah Salwa Nabila, Najla Khansa Alifah, Muhammad Tri Harsono. Terimakasih sudah bersama-sama menyelesaikan ini
11. Teman-teman seperjuangan Jurusan Matematika angkatan 2022.

Semoga skripsi ini dapat bermanfaat bagi kita semua. Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, sehingga penulis mengharapkan kritik dan saran yang membangun untuk menjadikan skripsi ini lebih baik lagi.

Bandar Lampung, 28 Januari 2026



Naufal Hafizd Prayoga

DAFTAR ISI

	Halaman
DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvi
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Tujuan Penelitian	5
1.3 Manfaat Penelitian	5
II TINJAUAN PUSTAKA	6
2.1 Eutrofikasi Danau	6
2.2 Upwelling Danau	7
2.3 Pemodelan Danau Satu Dimensi	7
2.3.1 Persamaan Satu Dimensi Difusi Panas	9
2.3.2 Atenuasi Radiasi Matahari	10
2.4 Artificial Neural Network (ANN)	11
2.4.1 Fungsi Aktivasi	12
2.4.2 Prinsip Matematis Backpropagation	14
2.5 Recurrent Neural Network (RNN)	15
2.6 Long Short-Term Memory (LSTM)	16
2.7 Convolutional Neural Network (CNN)	18
2.8 Random Forest (RF)	20
2.9 Physics-Informed Machine Learning	23
2.9.1 Physics-Informed Neural Network (PINN)	23
2.9.2 Mekanisme Perhitungan Turunan dengan Automatic Differentiation	24
2.9.3 Keunggulan PINN Dibandingkan Metode Numerik Tradisional	24
2.10 Aplikasi Deep Learning dalam Memprediksi Suhu Danau	26
2.11 Metrik Penelitian	27
III METODE PENELITIAN	30

3.1	Waktu dan Tempat Penelitian.....	30
3.2	Metode Penelitian.....	30
3.3	Studi Daerah Penelitian.....	31
3.3.1	Dinamika Stratifikasi dan Pencampuran Danau Batur	34
3.4	Data yang digunakan	35
3.4.1	Preprocessing Data.....	39
3.5	Desain Eksperimen dan Tahap Latih Model	44
3.5.1	Model yang Diuji	45
3.5.2	Model Hibrida PINN-LSTM	49
3.5.3	Implementasi Sumber Radiasi Matahari	50
3.5.4	Optimasi Hiperparameter	51
3.5.5	Pelatihan dan Evaluasi	52
3.6	Metrik Evaluasi	53
IV	HASIL DAN PEMBAHASAN	55
4.1	Hiperparameter Terbaik untuk Setiap Model	55
4.2	Tahapan Penelitian	58
4.2.1	Analisis Waktu Kinerja Model	59
4.2.2	Analisis Metrik Kinerja Umum.....	61
4.2.3	Analisis dan Pemilihan Parameter n_p dan n_f	67
4.2.4	Analisis Visual Model Terpilih: PINN-LSTM dan RF	73
4.2.5	Analisis Performa Random Forest dan PINN-LSTM	76
4.2.6	Implikasi Praktis untuk Sistem Peringatan Dini	77
4.2.7	Perbandingan dengan Penelitian Lainnya.....	79
V	KESIMPULAN DAN SARAN	81
5.1	Kesimpulan.....	81
5.2	Saran.....	82
	DAFTAR PUSTAKA	83
	LAMPIRAN	96

DAFTAR TABEL

	Halaman
2.1 Ringkasan Studi Terkait Aplikasi Machine Learning untuk Prediksi Suhu Air	28
3.1 Detail variabel yang digunakan sebagai prediktor dalam penelitian ini	40
3.2 Ringkasan Hyperparameter dari Literatur untuk Menjustifikasi Ruang Pencarian Optuna	46
3.3 Ringkasan Hyperparameter yang digunakan sebagai Ruang Pencarian Optuna	52
4.1 Hiperparameter Terpilih untuk Model CNN	56
4.2 Hiperparameter Terpilih untuk Model LSTM	56
4.3 Hiperparameter Terpilih untuk Model Random Forest	57
4.4 Hiperparameter Terpilih untuk Model PINN (MLP)	58
4.5 Hiperparameter Terpilih untuk Model PINN-LSTM	58
4.6 Perbandingan RMSE (°C) untuk Setiap Kombinasi Model dan Parameter	62
4.7 Perbandingan MAE (°C) untuk Setiap Kombinasi Model dan Parameter	62
4.8 Perbandingan NSE untuk Setiap Kombinasi Model dan Parameter ...	63
4.9 Perbandingan Ketidakkonsistenan Fisika (%) untuk Setiap Kombinasi Model dan Parameter	64
4.10 Perbandingan RMSE (°C) pada Periode Stratifikasi untuk Setiap Kombinasi Model dan Parameter	66
4.11 Perbandingan RMSE (°C) pada Periode Pencampuran untuk Setiap Kombinasi Model dan Parameter	66
4.12 Statistik Prediksi Suhu PINN-LSTM dan RF ($n_p = 72$, $n_f = 12$)	73
4.13 Perbandingan dengan Penelitian Prediksi Suhu Danau dan Perairan Lainnya	79

DAFTAR GAMBAR

	Halaman
1.1 Ribuan ikan nila atau ikan mujair mati di Danau Batur, Kintamani, Kabupaten Bangli, Bali (Posmerdeka, 2025)	2
2.1 Ilustrasi Eutrofikasi (Earthhow, 2017)	6
2.2 Ilustrasi Upwelling (Wiradana et al., 2022)	7
2.3 Struktur ANN	13
2.4 Contoh <i>Backpropagation</i>	15
2.5 Struktur RNN	16
2.6 Alur LSTM dalam satuan t	18
2.7 Struktur CNN	20
2.8 Struktur RF	22
2.9 Contoh Automatic Differentiation (AD)	25
3.1 Metode Penelitian	30
3.2 Laman SSH Mahameru BRIN HPC CPU	32
3.3 Peta danau Batur (Apple, 2025)	34
3.4 Struktur data dari berkas CSV	36
3.5 Siklus harian meteorologi	37
3.6 Data asli suhu Danau Batur diinterpolasi	38
3.7 Siklus harian suhu danau batur dibagi berdasarkan bulan	39
3.8 Seluruh Tahap <i>Preprocessing</i> Data	41
3.9 Batas $\pm 1.0^{\circ}\text{C}$ yang disebut sebagai fenomena mixing	42
3.10 Seluruh Tahap latih model	47
3.11 Abstraksi Arsitektur Model RF	47
3.12 Arsitektur Model LSTM	48
3.13 Arsitektur Model CNN	48
3.14 Arsitektur Model PINN	49
3.15 Arsitektur Model PINN-LSTM	50
3.16 Seluruh Tahap Evaluasi Data	54

4.1	Waktu pelatihan (detik) untuk Setiap Kombinasi Model	60
4.2	Waktu pengujian (detik) untuk Setiap Kombinasi Model	60
4.3	Persentase Ketidakkonsistenan Fisika Setiap Transisi Lapisan Kedalaman untuk Setiap Model	64
4.4	Persentase Ketidakkonsistenan Fisika Setiap Jam dalam Satu Hari untuk Setiap Model	65
4.5	Trade-off: Prediksi (RMSE) vs. Konsistensi Fisika pada Setiap Kombinasi Model	68
4.6	Analisis RMSE vs Horizon Prediksi pada Model Terpilih (PINN-LSTM dan RF)	69
4.7	Analisis Profil pada Model Terpilih dengan Variasi n_f	69
4.8	Analisis Profil pada Model Terpilih dengan Variasi n_p	70
4.9	Analisis RMSE vs Kedalaman pada Model Terpilih (PINN-LSTM dan RF)	71
4.10	Analisis Residual pada Model Terpilih (PINN-LSTM dan RF)	72
4.11	<i>Heatmap</i> Suhu Seluruh Periode Uji: Observasi vs PINN-LSTM vs RF ($n_p = 72, n_f = 12$)	74
4.12	<i>Heatmap</i> Suhu Diskret (Tanpa Interpolasi) Seluruh Periode Uji ($n_p =$ $72, n_f = 12$)	75
4.13	<i>Heatmap</i> Periode Transisi Pencampuran: Observasi vs PINN-LSTM vs RF ($n_p = 72, n_f = 12$)	75
4.14	<i>Heatmap</i> Periode Stratifikasi Stabil: Observasi vs PINN-LSTM vs RF ($n_p = 72, n_f = 12$)	76

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perubahan iklim global telah menjadi pendorong utama perubahan dalam ekosistem akuatik di seluruh dunia. Bukti empiris menunjukkan bahwa danau dan waduk tidak hanya mengalami pemanasan suhu permukaan secara global (Golub et al., 2022), tetapi juga peningkatan frekuensi dan durasi gelombang panas danau (Wang et al., 2023). Perubahan iklim juga memengaruhi proses hidrologis di daerah aliran sungai sehingga mengubah kuantitas dan kualitas air yang masuk ke danau. Kemudian, peningkatan suhu udara dan perubahan pola angin juga mengubah perilaku fisika danau, seperti memperkuat stratifikasi suhu dan mengurangi tutupan es yang pada akhirnya mengubah lingkungan ekosistem akuatik (Ozersky et al., 2025; Chowdhury et al., 2023).

Salah satu konsekuensi paling signifikan dari perubahan ini adalah perubahan pada dinamika pencampuran vertikal danau. Perubahan ini yang dapat memicu fenomena ekstrem seperti *upwelling* dan eutrofikasi. Secara limnologis, danau tropis dalam seperti Danau Batur memiliki karakteristik stratifikasi termal yang unik, dimana lapisan atas (*epilimnion*) yang hangat mengisolasi lapisan bawah (*hypolimnion*) yang dingin (Wetzel, 2001). Stratifikasi yang stabil ini bertindak sebagai "tutup" yang mencegah oksigen berdifusi ke dasar, menyebabkan kondisi tanpa oksigen (anoksia) dan penumpukan senyawa toksik seperti hidrogen sulfida (H_2S) dari dekomposisi organik (Boehrer and Schultze, 2008).

Bencana kematian ikan massal terjadi ketika "tutup" ini terbuka akibat mekanisme pencampuran (*overturn*) yang dipicu oleh angin kencang atau pendinginan permukaan yang drastis. Saat pencampuran terjadi, massa air anoksik dari dasar

naik ke permukaan (*upwelling*), menyebabkan penurunan kadar oksigen terlarut secara tiba-tiba yang mematikan bagi biota (Santoso et al., 2025a). Oleh karena itu, memprediksi suhu permukaan saja tidaklah cukup. Kunci mitigasi bencana ini terletak pada pemahaman struktur vertikal suhu air. Prediksi profil suhu satu dimensi (1D) menjadi krusial untuk mendeteksi seberapa kuat stabilitas stratifikasi dan kapan "titik kritis" pencampuran akan terjadi. Model berbasis data murni seringkali gagal menangkap dinamika stabilitas ini karena mengabaikan hukum kekekalan energi yang mengatur transisi tersebut, sehingga pendekatan berbasis fisika (*Physics-Informed*) menjadi solusi yang mendesak.

Fenomena eutrofikasi dan *upwelling* yang saling memperkuat ini sudah terjadi selama bertahun-tahun pada Danau Batur (Fajar, 2021). Surplus nutrisi dari aktivitas antropogenik memperparah kondisi anoksia di dasar, sementara *upwelling* mendaur ulang nutrisi tersebut kembali ke zona fotik. Konsekuensi fatal dari dinamika ini terlihat pada kasus kematian ikan massal tanggal 14 Juli 2025 (Wirawan, 2025; Wiradana et al., 2022) dengan estimasi kerugian 9 hingga 70 ton ikan. Gambar 1.1 menunjukkan dampak visual dari peristiwa tersebut.



Gambar 1.1 Ribuan ikan nila atau ikan mujair mati di Danau Batur, Kintamani, Kabupaten Bangli, Bali (Posmerdeka, 2025)

Berbagai studi telah dilakukan untuk menjawab permasalahan danau di Indonesia, terutama Danau Batur. Berbagai studi tersebut meliputi aspek pencegahan, pengawasan atau monitoring, dan perbaikan atau korektif. Contoh studi untuk meningkatkan pengawasan, yaitu pengembangan *Geographic Information System* (GIS) sebagai alat bantu pengambilan keputusan dalam pengelolaan KJA yang berkelanjutan di Danau Batur (Yustiati and Lusia, 2025) dan *Decision Support System* (DSS) (Pen, 2024) untuk mengendalikan pencemaran lingkungan di Danau Batur. Studi lainnya adalah pengembangan model numerik berbasis web untuk mensimulasikan variabel-variabel utama terkait kualitas air menggunakan *General Lake Model* (GLM) dengan modul *Aquatic Ecological Dynamic* (AED) (Sunaryani et al., 2023). Dari berbagai sistem yang telah dikembangkan tersebut, kemampuan untuk memprediksi menjadi penting agar sistem tidak hanya memberikan laporan kejadian tetapi juga memberikan masukan pengambil keputusan untuk mencegah kejadian atau mengurangi dampak dari suatu kejadian.

Di sisi lain, *machine learning*, terutama deep learning, telah berkembang dan terbukti mampu memprediksi berbagai data *time-series* dengan kemampuan komputasi yang efisien sehingga menjadi alternatif metode atau model konvensional. Berbagai model *machine learning* yang digunakan dalam studi danau antara lain *Long Short-Term Memory* (LSTM) (Qiu et al., 2021; Wang et al., 2022), *Convolutional Neural Networks* (CNN) (Zhu et al., 2024; Barzegar et al., 2020), *Random Forest* (RF) (Heddami et al., 2020; Dyba et al., 2022), *Gated Recurrent Unit* (GRU) (Hao et al., 2023b; He and Yang, 2025), *Physics Informed Machine Learning* (PIML) (Lawal et al., 2022; He and Yang, 2025), dan *Physics Guided Machine Learning* (Daw et al., 2021; Jia et al., 2020). Adapun variabel yang dipredikasi dalam berbagai studi tersebut adalah suhu danau yang merupakan indikator krusial dalam monitoring kejadian *upwelling* (Kayastha et al., 2023; Willard et al., 2022). Studi-studi tersebut dilakukan pada berbagai danau, waduk, dan sungai di seluruh dunia, yang mencakup studi kasus pada Danau Qiandaohu (He and Yang, 2025), Waduk Nuozhadu (Wang et al., 2022), dan Waduk Xiaowan (Wang et al., 2022) di Tiongkok; Danau Mendota (Jia et al., 2020; Read et al., 2019; Daw et al., 2021), Danau Mille Lacs (Daw et al., 2021), Danau Sparkling (Read et al., 2019), dan Danau-Danau Besar (*the*

Great Lakes) (Kayastha et al., 2023) di Amerika Serikat; Danau Kasumigaura (Zhu et al., 2024) di Jepang; Waduk Sau (Mercado-Bettín et al., 2021) di Spanyol, Waduk Wupper (Mercado-Bettín et al., 2021) di Jerman, Waduk Mt. Bold di Australia, Danau Vansjø (Mercado-Bettín et al., 2021) di Norwegia; serta berbagai danau di Polandia (Zhu et al., 2020), dan Yunani (Barzegar et al., 2020).

Dari berbagai studi yang telah dilakukan terlihat bahwa pengembangan model prediksi berbasis *machine learning* untuk danau di Indonesia belum banyak dilakukan. Berdasarkan aspek limnologi, danau di Indonesia sebagai danau tropis memiliki perbedaan fundamental dengan danau bersuhu dingin di wilayah subtropis dalam dinamika pemisahan (stratifikasi) dan pencampuran (*mixing*) air di dalam danau (Pu et al., 2025). Studi terbaru oleh Santoso et al. (2025a) menunjukkan bahwa Danau Batur memiliki siklus polimiktik (*polymictic*) dengan empat periode musiman: stratifikasi awal (Oktober-Desember), pencampuran (Desember-Maret), stratifikasi kuat (Maret-Mei), dan pencampuran sering (Mei-Oktober). Pencampuran vertikal ini dapat membawa air anoksik yang kaya sulfida dari dasar danau ke permukaan, mengakibatkan kematian ikan massal (Santoso et al., 2025a). Terlebih lagi Danau Batur yang memiliki keunikan berupa tidak adanya aliran masuk (inlet) dan aliran keluar (outlet), serta tinggi belerang. Oleh karena itu, dibutuhkan sistem peringatan dini yang mampu memprediksi kejadian pencampuran ini sebelum terjadi. Salah satu metode yang dapat menjawab kebutuhan ini adalah *Physics-Informed Neural Networks* (PINN). PINN merupakan sebuah arsitektur *deep learning* yang menggunakan pendekatan hibrida antara model berbasis data dan berbasis hukum-hukum fisika. Dengan berbasis PINN, model yang dikembangkan diharapkan tidak hanya mengenali pola pada data, tetapi juga sesuai dengan siklus biogeokimia pada danau sehingga dapat diterima oleh para limnologis (He and Yang, 2025).

Namun, PINN tidak dapat digunakan untuk waktu jangka panjang (He and Yang, 2025). maka dari itu terdapat model baru yang disebut PINN-LSTM, yang menggabungkan PINN dengan kemampuan LSTM dalam memodelkan data sekuensial. Arsitektur hibrida ini terinspirasi oleh keberhasilan penerapan PINN-LSTM untuk peramalan gelombang laut (Lawal et al., 2025), namun dengan

modifikasi untuk mensimulasikan dinamika suhu danau. PINN-LSTM ini akan dibandingkan dengan model *machine learning* lainnya yang sudah berhasil dalam memprediksi suhu danau, yaitu LSTM, CNN, dan RF. Model tersebut akan digunakan untuk memprediksi profil suhu air di Danau Batur.

1.2 Tujuan Penelitian

Berdasarkan latar belakang dan perumusan masalah pada bagian sebelumnya, tujuan penelitian yang akan dilakukan adalah sebagai berikut.

1. Mengembangkan arsitektur *deep learning* yang menggabungkan hukum fisika dalam proses pelatihannya untuk memprediksi secara *time-series* profil suhu di Danau Batur dengan akurat dan cepat.
2. Mengevaluasi kinerja arsitektur yang akan dikembangkan dan membandingkan hasil prediksi model dengan berbagai arsitektur model yang digunakan dalam penelitian-penelitian sebelumnya, seperti RF, LSTM, CNN, dan PINN.

1.3 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut.

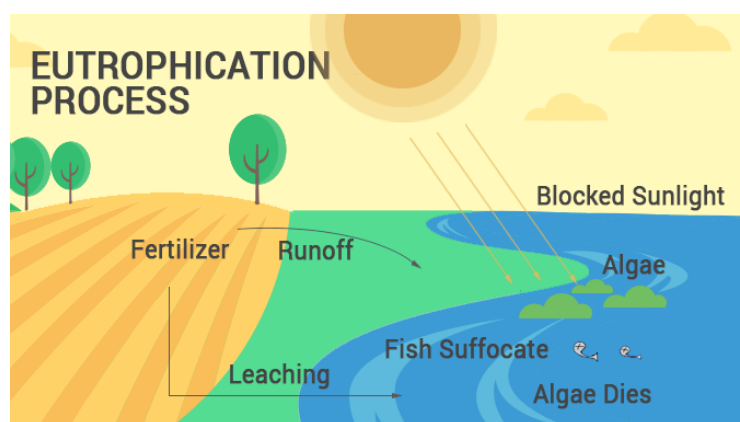
1. Memberikan pemahaman yang lebih dalam tentang kekuatan dan kelemahan berbagai arsitektur *machine learning* untuk pemodelan limnologi, khususnya prediksi profil suhu danau.
2. Menghasilkan alternatif metode untuk memprediksi suhu danau sebagai bagian dari sistem monitoring untuk fenomena *upwelling* dan kematian ikan.
3. Memberikan rekomendasi berbasis bukti mengenai pemilihan model untuk para peneliti dan praktisi di bidang limnologi dan pemodelan lingkungan.

BAB II

TINJAUAN PUSTAKA

2.1 Eutrofikasi Danau

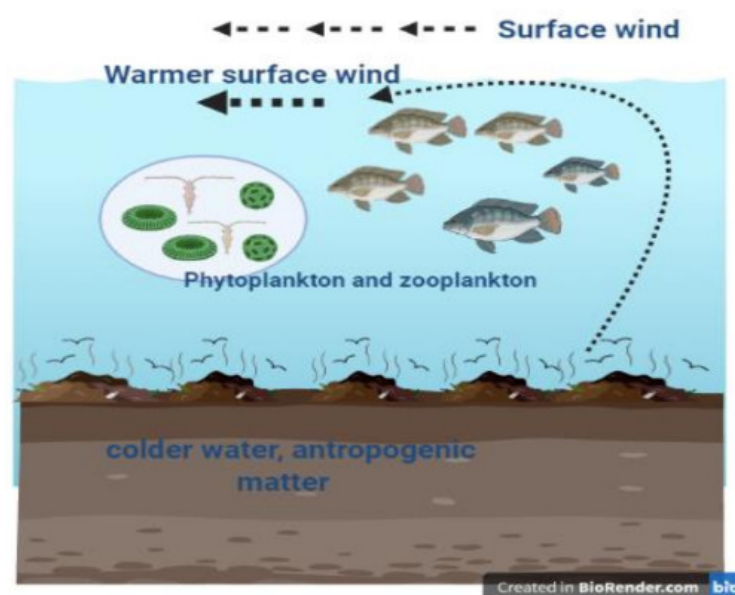
Menurut Akinnowo (2023), Eutrofikasi (*eutrophication*) berasal Yunani dari kata *eutrophic* yang memiliki arti “kaya-nutrisi”. Eutrofikasi didefinisikan sebagai kondisi perairan yang memiliki konsentrasi nutrisi yang tinggi, termasuk nitrogen dan fosfor, yang menyebabkan terjadinya ledakan alga, sehingga menurunkan kualitas air dalam ekosistem perairan. Eutrofikasi adalah proses dapat didasari atas faktor alami maupun antropogenik (faktor manusia). Faktor antropogenik dapat berupa penebaran alami suatu badan air hingga perubahan status eutrofik (Sonarghare et al., 2020). Namun, pembuangan limbah serta beban nutrisi terbatas seperti nitrogen dan fosfor yang berasal dari aktivitas manusia telah mempercepat laju dan skala eutrofikasi. Sehingga dengan konsekuensi yang merugikan air danau sumber air minum, kehidupan perairan, serta rekreasi (Zhang et al., 2023). Ilustrasi eutrofikasi terlihat pada Gambar 2.1



Gambar 2.1 Ilustrasi Eutrofikasi (Earthhow, 2017)

2.2 Upwelling Danau

Menurut definisi yang diberikan oleh Merriam-Webster (2025), *Upwelling* merupakan proses pergerakan air dari dasar laut ke atas menuju permukaan laut. Pergerakan ini membawa air yang dingin, kaya nutrisi, namun sedikit oksigen terlarut ke arah lepas permukaan. Fenomena *upwelling* terjadi akibat perbedaan distribusi panas di lapisan-lapisan danau. Angin di udara dapat menghasilkan energi kinetik besar di perairan, menyebabkan pergerakan massa air dan pencampuran di danau secara horizontal dan vertikal. Hal ini terjadi terutama untuk danau dengan kondisi stratifikasi (Shintani et al., 2010; Li et al., 2021). Gaya Coriolis dan geometri danau menjadi pengaruh utama dalam bentuk pola *upwelling* di danau (Roberts et al., 2021). Ilustrasi fenomena *upwelling* tersaji dalam Gambar 2.2.



Gambar 2.2 Ilustrasi Upwelling (Wiradana et al., 2022)

2.3 Pemodelan Danau Satu Dimensi

Salah satu penelitian penting adalah penelitian yang menerapkan GLM-AED dalam simulasi perairan Danau Batur. GLM merupakan salah satu metode numerik yang

dikembangkan untuk melakukan simulasi terhadap berbagai hal di danau (Hipsey et al., 2019). Contoh simulasi yang dapat digunakan GLM seperti stratifikasi suhu vertikal, pencampuran suhu, dan aliran masuk atau keluar air, serta pemanasan dan pendinginan permukaan danau. Luaran dari model GLM beragam, seperti suhu danau, oksigen terlarut, klorofil-a, fosfat, amonia, nitrit, nitrogen total, fosfor total, dan lain lain (Sunaryani et al., 2023).

Salah satu luaran yang utama dalam GLM adalah suhu danau. Hal ini dikarenakan suhu danau merupakan salah satu variabel terpenting dalam memahami proses pencampuran dan biogeokimia. Suhu memiliki korelasi terhadap jumlah oksigen yang terlarut dalam danau. Lapisan permukaan bertindak sebagai sumber oksigen untuk danau dan lapisan dasar tidak dapat memproduksi oksigen sehingga menghabiskan oksigen melalui mineralisasi bahan organik dan sedimen (Müller et al., 2012). Kemampuan untuk memprediksi suhu air danau secara akurat adalah langkah pertama untuk memahami dan mengantisipasi kejadian ekologis yang merugikan seperti kematian ikan massal yang terjadi pada Danau Batur.

Model fisika danau yang digunakan di GLM dalam memprediksi suhu danau adalah model satu dimensi (1D) vertikal. Model-model ini didasarkan pada prinsip konservasi energi dan melakukan penyederhanaan dalam dinamika danau dengan mengasumsikan bahwa variasi terjadi terutama pada arah vertikal (kedalaman), sementara kondisi horizontal dianggap seragam (Hipsey et al., 2019). Hal ini dikarenakan kompleksitas dari koefisien yang ada dan adanya sumber panas eksternal.

GLM telah digunakan secara luas dan menunjukkan kinerja yang baik di banyak danau, terutama di wilayah subtropis (Hipsey et al., 2019; Feldbauer et al., 2025). Namun, model-model ini seringkali membutuhkan kalibrasi yang ekstensif dan biaya komputasi yang tinggi. Kalibrasi yang GLM butuhkan seperti ketebalan lapisan danau, koefisien pertukaran panas, efisiensi pencampuran suhu pada danau, turbulensi danau, dan lain lain (Hipsey et al., 2019). GLM juga dapat mengalami penurunan akurasi pada danau-danau tropis yang kompleks (Guo et al., 2023; Pu et al., 2025) seperti Danau Batur dan Danau Towuti dikarenakan minimnya

pengumpulan observasi di daerah tersebut.

2.3.1 Persamaan Satu Dimensi Difusi Panas

Perubahan suhu air danau (T) terhadap waktu (t) dan kedalaman (z) diatur oleh persamaan satu dimensi difusi panas. Persamaan ini digunakan untuk menjaga keseimbangan antara penyimpanan panas dengan transpor panas vertikal dan sumber panas dari radiasi matahari (He and Yang, 2025; Guo et al., 2020). Persamaan ini umumnya tidak memiliki solusi analitis dan diselesaikan menggunakan metode numerik (LeVeque, 2007).

$$c_w \frac{\partial T}{\partial t} = \frac{\partial}{\partial z} \left(\tau \frac{\partial T}{\partial z} \right) - \frac{d\phi}{dz} \quad (2.3.1)$$

dimana:

- T adalah suhu air danau (K)
- t adalah waktu (s)
- z adalah kedalaman (m).
- c_w adalah kapasitas panas volumetrik air ($\text{J m}^{-3} \text{K}^{-1}$).
- τ adalah difusivitas panas total ($\text{m}^2 \text{s}^{-1}$), yang mencakup difusivitas molekuler dan difusivitas turbulen (*eddy diffusivity*) yang terutama disebabkan oleh angin.
- ϕ adalah fluks radiasi matahari (W m^{-2}) pada kedalaman z .

Untuk menggunakan persamaan 2.3.1, terdapat asumsi yang diperlukan, yaitu asumsi homogenitas horizontal yang menyatakan bahwa untuk setiap lapisan kedalaman, suhu air diasumsikan seragam secara horizontal. Dengan kata lain, suhu pada kedalaman z tertentu tidak bervariasi terhadap koordinat horizontal (x, y), sehingga $T = T(z, t)$. Asumsi ini digunakan karena Danau Batur memiliki ukuran

danau relatif kecil (≈ 7 km), di mana pencampuran horizontal terjadi lebih cepat dibandingkan dengan stratifikasi termal secara vertikal (Imberger and Csanady, 2015; Santoso et al., 2025a).

2.3.2 Atenuasi Radiasi Matahari

Fluks radiasi matahari berkurang secara eksponensial seiring dengan bertambahnya kedalaman. Fenomena ini dijelaskan oleh Hukum Beer-Lambert (Swinehart, 1962), dimana sebagian dari radiasi gelombang pendek diserap di permukaan dan sisanya menembus ke dalam lapisan air.

$$\phi(z) = (1 - \beta)I_{sol}e^{-\eta z} \quad (2.3.2)$$

dimana:

- I_{sol} adalah total radiasi matahari gelombang pendek yang mencapai permukaan air (W m^{-2}).
- β adalah fraksi radiasi yang diserap di lapisan permukaan.
- η adalah koefisien atenuasi cahaya (m^{-1}), yang menentukan tingkat kejernihan air.

Berdasarkan Persamaan 2.3.2, fluks radiasi matahari pada kedalaman yang berbeda dapat dihitung. Hal ini menjelaskan mengapa lapisan permukaan memiliki variabilitas suhu yang tinggi sedangkan lapisan dalam memiliki suhu yang relatif stabil sepanjang tahun. Sebagai contoh, dengan asumsi $\beta = 0.4$ dan $\eta = 0.15 \text{ m}^{-1}$, diperoleh

- $\phi(0) = 0.6 \times I_{sol}$ (Kedalaman 0 meter)
- $\phi(10) = 0.6 \times I_{sol} \times e^{-1.5} \approx 0.134 \times I_{sol}$ (Kedalaman 10 meter)
- $\phi(20) = 0.6 \times I_{sol} \times e^{-3.0} \approx 0.030 \times I_{sol}$ (Kedalaman 20 meter)
- $\phi(55) \approx 0$ (Kedalaman 55 meter)

2.4 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) yang terinspirasi oleh struktur dan fungsi jaringan saraf biologis di otak manusia (McCulloch and Pitts, 1943). Struktur ANN disajikan dalam Gambar 2.3. ANN terdiri dari unit-unit pemrosesan sederhana yang disebut *neuron*, yang diorganisir dalam beberapa lapisan (*layers*). Lapisan-lapisan ini umumnya terbagi menjadi tiga jenis: lapisan masukan (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layers*), dan lapisan keluaran (*output layer*) (Rosenblatt, 1958).

Secara matematis, operasi dalam satu neuron dapat direpresentasikan sebagai berikut:

$$y = \sigma \left(\sum_{i=1}^n w_i x_i + b \right) \quad (2.4.3)$$

di mana:

- y adalah output dari neuron.
- x_i adalah input ke- i .
- w_i adalah bobot yang terkait dengan input ke- i .
- b adalah bias.
- n adalah jumlah input.
- σ adalah fungsi aktivasi (misalnya, fungsi sigmoid, ReLU, atau tanh) yang memperkenalkan non-linearitas ke dalam model.

Untuk keseluruhan lapisan, operasi ini dapat diekspresikan dalam bentuk matriks:

$$\mathbf{y} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (2.4.4)$$

2.4.1 Fungsi Aktivasi

Fungsi aktivasi merupakan komponen krusial dalam sebuah *neuron* yang berfungsi untuk memperkenalkan non-linearitas ke dalam model. Tanpa fungsi aktivasi non-linear, sebuah jaringan saraf, tidak peduli berapa banyak lapisannya, secara matematis akan setara dengan model regresi linear sederhana (Goodfellow et al., 2016).

Terdapat beberapa fungsi aktivasi yang umum digunakan, di antaranya (Goodfellow et al., 2016):

- Fungsi Sigmoid

Fungsi ini memetakan nilai input ke dalam rentang antara 0 dan 1. Fungsi ini sering digunakan pada lapisan keluaran untuk klasifikasi biner. Namun, fungsi ini dapat terjadi masalah *vanishing gradient*, dimana gradien menjadi sangat kecil untuk nilai input yang sangat besar atau sangat kecil, sehingga memperlambat proses pelatihan.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.4.5)$$

- Fungsi Tangen Hiperbolik (Tanh)

Mirip dengan sigmoid, tetapi memetakan input ke rentang antara -1 dan 1. Karena luarannya berpusat di nol (*zero-centered*), Tanh konvergen lebih cepat daripada Sigmoid, Namun tetap rentan terhadap masalah *vanishing gradient*.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.4.6)$$

- *Rectified Linear Unit* (ReLU)

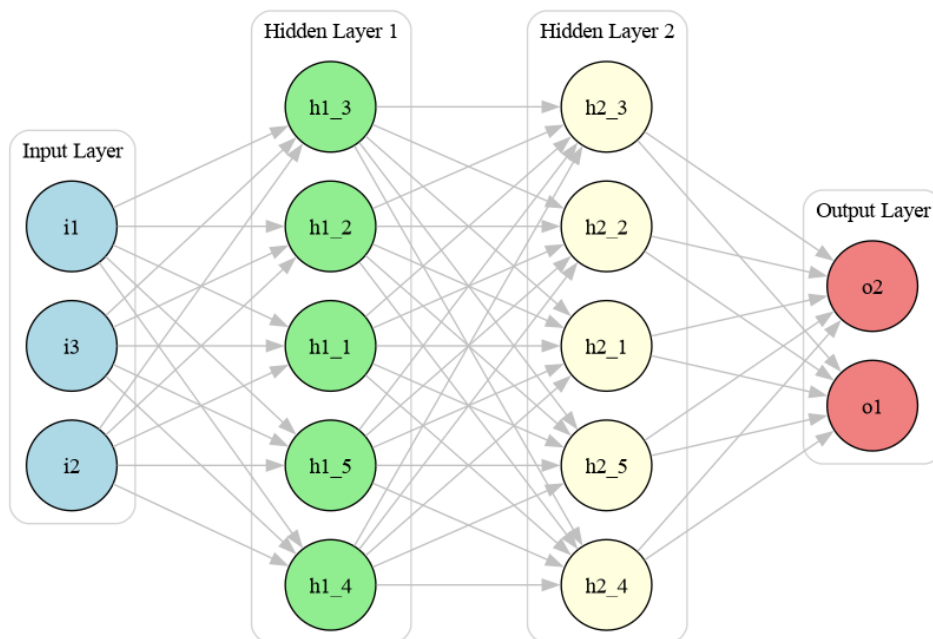
Fungsi ini mengembalikan nilai input jika positif, dan nol jika negatif. ReLU merupakan fungsi aktivasi paling populer untuk *hidden layer* karena efisiensi komputasi dan kemampuannya untuk mengatasi masalah *vanishing gradient*

pada arah positif.

$$\text{ReLU}(x) = \max(0, x) \quad (2.4.7)$$

Pemilihan fungsi aktivasi yang tepat merupakan salah satu aspek penting dalam desain arsitektur jaringan saraf, karena dapat memengaruhi kecepatan konvergensi dan performa akhir model (Goodfellow et al., 2016).

Setiap koneksi antar neuron memiliki bobot (*weight*) yang menentukan seberapa besar nilai yang dikeluarkan. Selama proses pelatihan, model menyesuaikan bobot-bobot ini untuk meminimalkan perbedaan antara output yang diprediksi dan output yang sebenarnya, sebuah proses yang dikenal sebagai *backpropagation* (Rumelhart et al., 1986).



Gambar 2.3 Struktur ANN

Algoritma *backpropagation* merupakan algoritma penting yang digunakan untuk hampir seluruh model DL. Backpropagation juga digunakan dalam suatu teknik yang disebut *Automatic Differentiation* (AD). Mekanisme ini memungkinkan jaringan untuk belajar dari data dengan menyesuaikan parameter internalnya secara iteratif untuk meminimalkan kesalahan prediksi.

2.4.2 Prinsip Matematis Backpropagation

Backpropagation, yang merupakan singkatan dari “*backward propagation of errors*”, adalah metode untuk menghitung gradien dari *loss function* terhadap seluruh bobot dan bias dalam jaringan saraf secara efisien (Rumelhart et al., 1986). Proses ini pada dasarnya adalah pengaplikasian dari aturan rantai (*chain rule*) yang dibagi menjadi dua tahap:

1. *Forward Pass*.

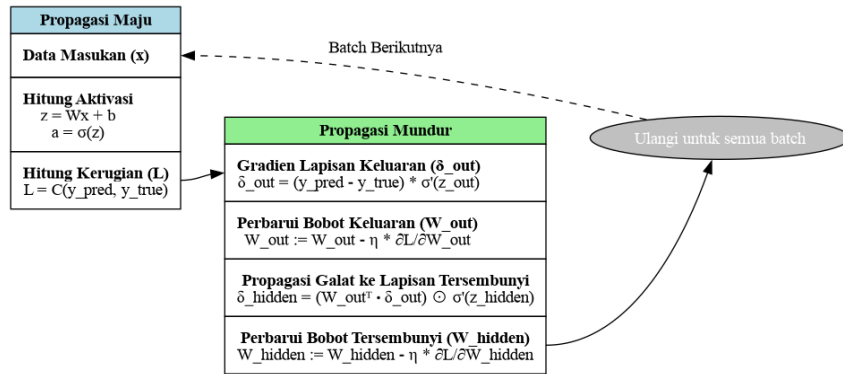
Input data dimasukkan ke dalam jaringan dari beberapa lapisan. Setiap lapisan akan dihitung *output* berdasarkan bobot, bias, dan fungsi aktivasi. *Output* akhir dari jaringan dibandingkan dengan nilai sebenarnya untuk menghitung nilai kesalahan total, yang diukur oleh sebuah *loss function*.

2. *Backward Pass*

Hasil dari *loss function* akan dihitung gradien terhadap *output* lapisan tersebut. Kemudian, dengan menggunakan aturan rantai, gradien ini disebarkan mundur (*propagated backward*) melalui jaringan setiap lapisan. Gradien dari *loss* terhadap bobot dan bias lapisan tersebut dihitung. Gradien ini kemudian digunakan oleh algoritma optimisasi untuk memperbarui nilai bobot dan bias ke arah yang akan mengurangi kesalahan.

Dengan mengulangi kedua tahap ini berulang kali untuk seluruh data pelatihan, jaringan secara bertahap “belajar” untuk menghasilkan *output* yang semakin mendekati nilai target. Contoh matematis tersaji pada Gambar 2.4.

Salah satu tantangan signifikan dalam melakukan *machine learning* adalah masalah vanishing gradient. Selama fase *backward pass*, gradien dari *loss function* dihitung terhadap setiap bobot dengan menggunakan aturan rantai. Jika jaringan memiliki banyak lapisan, proses ini melibatkan perkalian berulang dari banyak turunan kecil. Apabila nilai gradien secara konsisten kurang dari 1, hasil perkaliannya akan menurun secara eksponensial mendekati nol saat mencapai lapisan-lapisan awal jaringan. Akibatnya, bobotnya tidak dapat diperbaharui dengan nilai baru (Bengio

Gambar 2.4 Contoh *Backpropagation*

et al., 1994; Goodfellow et al., 2016).

2.5 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) merupakan pengembangan dari ANN yang dirancang khusus untuk menangani data sekuensial seperti data deret waktu atau teks (Elman, 1990). Berbeda dengan ANN, RNN memiliki koneksi berulang yang memungkinkan informasi untuk bertahan dari satu langkah waktu ke langkah waktu berikutnya. Contoh Struktur RNN disajikan pada Gambar 2.5. Mekanisme ini diimplementasikan melalui penggunaan *hidden state* (h_t), yang pada setiap langkah waktu t tidak hanya menerima masukan saat ini (x_t) tetapi juga menerima *hidden state* dari langkah waktu sebelumnya (h_{t-1}) (Goodfellow et al., 2016).

Secara matematis, hubungan ini diekspresikan sebagai berikut:

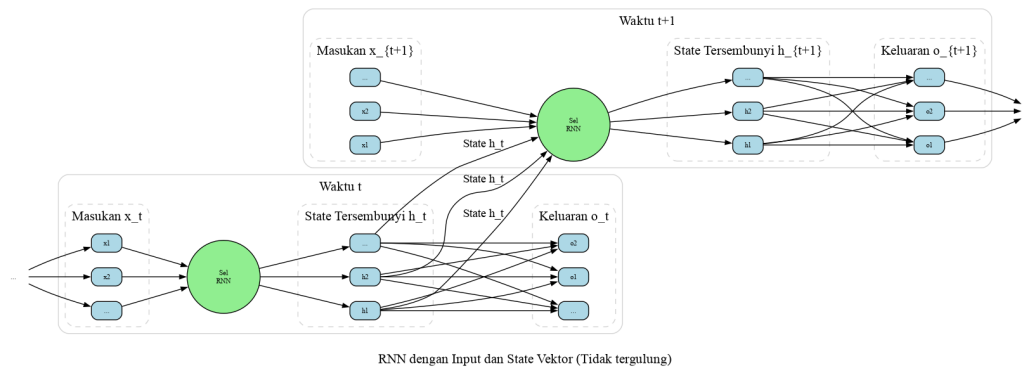
$$h_t = \sigma_h(W_{hh}h_{t-1} + W_{hx}x_t + b_h) \quad (2.5.8)$$

$$y_t = \sigma_y(W_{hy}h_t + b_y) \quad (2.5.9)$$

di mana:

- h_t adalah *hidden state* pada langkah waktu t .
- x_t adalah input pada langkah waktu t .

- h_{t-1} adalah *hidden state* dari langkah waktu sebelumnya.
- y_t adalah output pada langkah waktu t .
- W_{hh}, W_{xh}, W_{hy} adalah matriks bobot untuk koneksi *hidden-to-hidden*, *input-to-hidden*, dan *hidden-to-output*.
- b_h dan b_y adalah vektor bias.
- σ_h dan σ_y adalah fungsi aktivasi (misalnya, *tanh* atau *ReLU* untuk σ_h).



Gambar 2.5 Struktur RNN

Kemampuan untuk mengingat informasi masa lalu menjadikan RNN efektif dalam memahami konteks sekuensial. Namun, arsitektur RNN dasar memiliki kelemahan signifikan yang dikenal sebagai masalah *vanishing gradient*, di mana gradien yang digunakan untuk memperbarui bobot jaringan menjadi sangat kecil selama proses *backpropagation* pada sekuens panjang. Hal ini membuat RNN kesulitan untuk mempelajari dependensi jangka panjang dalam data (Bengio et al., 1994).

2.6 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) merupakan model yang dikembangkan oleh Hochreiter and Schmidhuber (1997) dan diteruskan oleh Gers et al. (2000). LSTM dikembangkan karena RNN mengalami masalah *vanishing gradient*. LSTM memperkenalkan *memory cell* dan dua *gate cell* untuk mengatur aliran informasi

(Hochreiter and Schmidhuber, 1997). Penelitian Gers et al. (2000) menambahkan satu jenis *gate cell* sehingga totalnya untuk LSTM terdapat tiga *gate cell*.

Tiga jenis *gate cell* tersebut dinamakan *forget gate*, *input gate*, dan *output gate*. *Forget gate* menentukan bagian mana dari keadaan sel sebelumnya yang akan dipertahankan atau dibuang. *Input gate* memperbarui keadaan sel dengan informasi baru yang relevan dari masukan. *Output gate* menghasilkan *hidden state*, yang digunakan untuk prediksi atau diteruskan ke langkah waktu berikutnya (Hochreiter and Schmidhuber, 1997).

Secara matematik, operasi di LSTM dalam satuan waktu t dapat dijelaskan sebagai berikut:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.6.10)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.6.11)$$

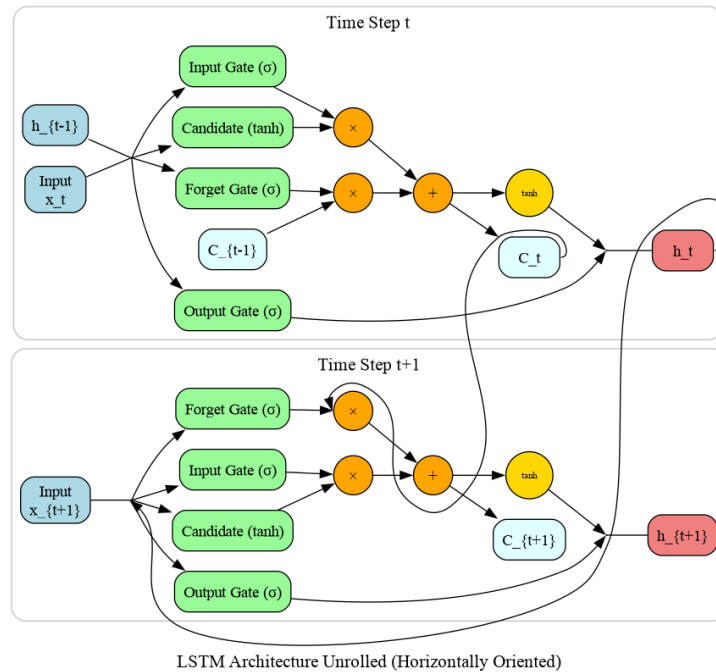
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2.6.12)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (2.6.13)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.6.14)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (2.6.15)$$

Persamaan 2.6.10, 2.6.11, 2.6.12, 2.6.12, 2.6.13, 2.6.14, 2.6.15 berturut turut adalah *forget gate*, *input gate*, *candidate cell state*, *cell state update*, *output gate*, dan *hidden gate*. Alur operasi persamaan LSTM ditampilkan dalam Gambar 2.6. Untuk menjalankan satu unit LSTM pada langkah waktu t , *input* x_t , *hidden gate* sebelumnya h_{t-1} , dan *cell state* sebelumnya C_{t-1} adalah sesuai dengan menghitung dengan urutan yang disebutkan sebelumnya (Hochreiter and Schmidhuber, 1997).



Gambar 2.6 Alur LSTM dalam satuan t

2.7 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah model dari *machine learning* yang dirancang untuk memproses data yang memiliki struktur seperti *grid*, contohnya adalah data citra (2D grid) (LeCun et al., 1998). Arsitektur CNN terinspirasi oleh korteks visual biologis, di mana neuron merespons rangsangan hanya di area terbatas yang dikenal sebagai *Receptive Field*.

Keunggulan utama CNN dibandingkan dengan ANN biasa terletak pada *parameter sharing* dan *translation invariance*. Dalam ANN, setiap neuron di satu lapisan terhubung ke setiap neuron di lapisan berikutnya, masing-masing dengan bobotnya sendiri. Sedangkan CNN, terdapat *filter* atau *kernel* yang sama—dengan satu set parameter yang sama untuk digunakan di seluruh input. Sehingga model lebih efisien dan mengurangi risiko *overfitting*. Karena filter yang sama diterapkan di seluruh input, CNN dapat mendeteksi pola yang sama di mana pun pola itu muncul dalam data. Misalnya, dalam data deret waktu, pola kenaikan suhu yang sama dapat diidentifikasi baik di awal maupun di akhir.

Operasi fundamental dalam CNN adalah konvolusi. Untuk data sekuensial (1D), operasi konvolusi pada suatu titik t dari sekuens input x dengan sebuah filter (kernel) w dapat ditulis sebagai:

$$y(t) = (x * w)(t) = \sum_{k=1}^K w_k \cdot x_{t-k+1} \quad (2.7.16)$$

di mana:

- $y(t)$ adalah elemen output pada posisi t .
- x adalah sekuens input.
- w adalah filter (kernel) dengan panjang K .
- k adalah indeks dari filter.

Filter ini digeser sepanjang sekuens input untuk menghasilkan *feature map*, Sehingga dapat mengidentifikasi pola-pola tertentu dalam data.

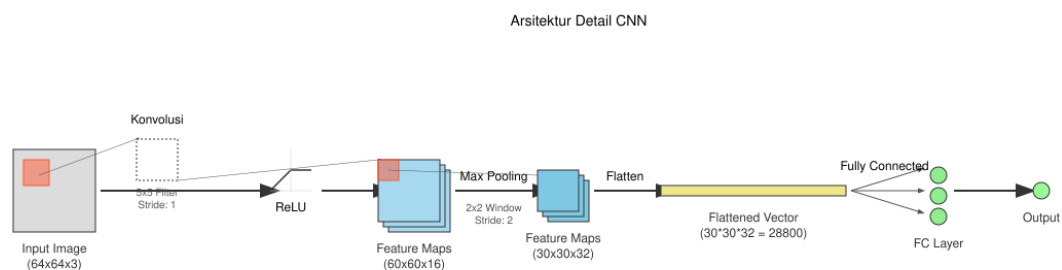
Arsitektur CNN terdiri dari beberapa jenis lapisan utama (LeCun et al., 1998):

- Lapisan Konvolusi (*Convolutional Layer*)
Lapisan ini menerapkan sejumlah *filter* ke input. Setiap *filter* secara independen mempelajari untuk mendeteksi fitur tertentu. Output dari operasi konvolusi kemudian biasanya dilewatkan melalui fungsi aktivasi non-linear, seperti ReLU (*Rectified Linear Unit*), untuk memperkenalkan non-linearitas. Dua parameter penting di lapisan ini adalah: *Stride* dan *Padding*. *Stride* menentukan berapa banyak langkah *filter* bergeser pada setiap konvolusi. *Stride* yang lebih besar akan menghasilkan *feature map* yang lebih kecil. *Padding* menambahkan nilai (biasanya nol) di sekitar batas input untuk memungkinkan *filter* beroperasi di tepi data dan mengontrol dimensi luaran.
- Lapisan Pooling (*Pooling Layer*)
Lapisan ini bertujuan untuk mengurangi dimensi spasial (ukuran) dari representasi, sehingga mengurangi jumlah parameter dan komputasi dalam

jaringan. Hal ini juga membantu membuat representasi fitur menjadi lebih kuat terhadap variasi kecil pada posisi fitur. Jenis pooling yang paling umum adalah *Max Pooling* dan *Average Pooling*. *Max Pooling* mengambil nilai maksimum dari setiap daerah di *feature map* dan *Average Pooling* menghitung nilai rata-rata dari setiap daerah di *feature map*.

- Lapisan Terhubung Penuh (*Fully-Connected Layer*)

Setelah beberapa lapisan konvolusi dan *pooling*, fitur-fitur tingkat tinggi yang telah dipelajari kemudian diratakan (*flattened*) menjadi vektor satu dimensi. Vektor ini kemudian menjadi masukan untuk satu atau lebih lapisan *fully-connected*, yang pada dasarnya adalah model ANN. Lapisan-lapisan ini digunakan untuk menggabungkan fitur-fitur yang diekstraksi untuk melakukan seperti klasifikasi atau regresi.



Gambar 2.7 Struktur CNN

2.8 Random Forest (RF)

Random Forest (RF) adalah metode pembelajaran *ensemble* yang dapat digunakan untuk tugas klasifikasi dan regresi (Breiman, 2001). Algoritma ini bekerja dengan membangun sejumlah besar *decision trees* pada saat pelatihan. Prediksi akhir dari model RF adalah rata-rata dari prediksi yang dihasilkan oleh setiap pohon individu. Ilustrasi arsitektur RF tersaji pada Gambar 2.8

Secara matematis, untuk tugas regresi, prediksi dari model Random Forest dengan

B pohon keputusan adalah rata-rata dari prediksi setiap pohon individu:

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (2.8.17)$$

di mana:

- $\hat{f}(x)$ adalah prediksi akhir untuk input x .
- B adalah jumlah total pohon dalam hutan.
- $T_b(x)$ adalah prediksi yang dihasilkan oleh pohon ke- b .

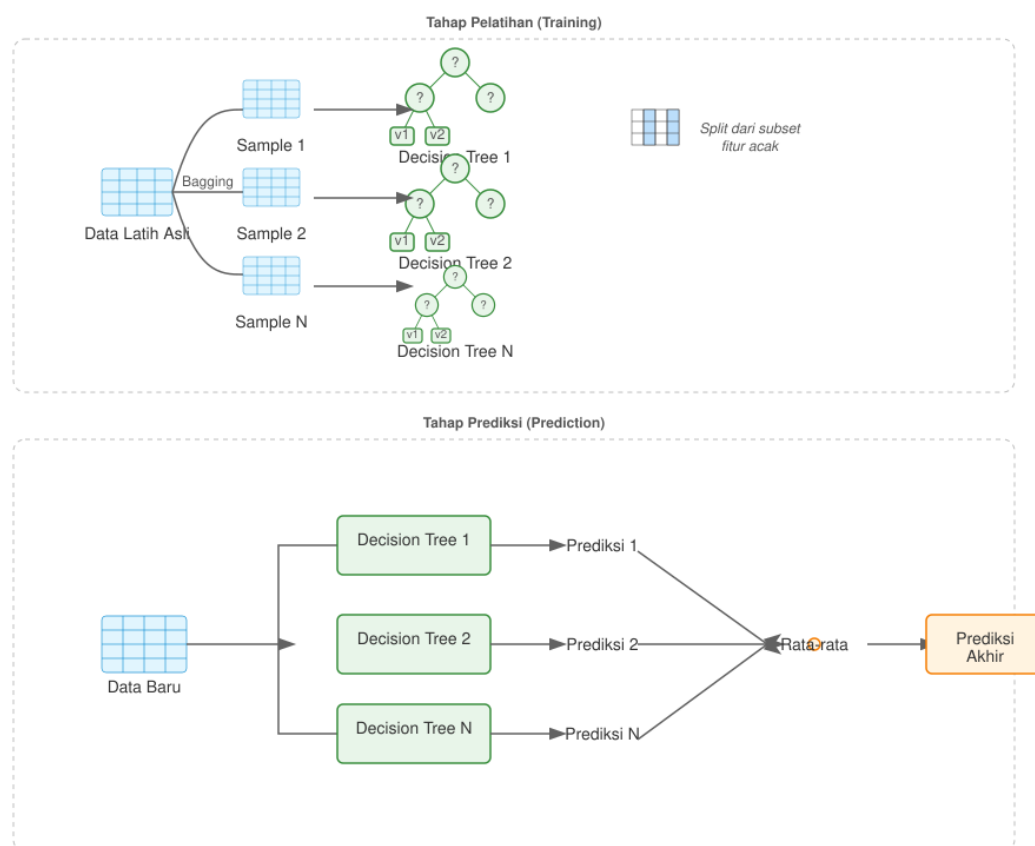
Untuk tugas klasifikasi, prediksi akhir ditentukan oleh suara mayoritas (*majority vote*) dari semua pohon.

Kekuatan RF terletak pada dua aspek utama dari “keacakan” yang diperkenalkan selama proses pelatihan:

- *Bagging (Bootstrap Aggregating)*
Setiap pohon dalam hutan dibangun dari sampel acak yang diambil dari set data pelatihan dengan penggantian (*bootstrap sample*). Setiap pohon yang dilatih pada subset data sedikit berbeda, sehingga membantu mengurangi varians dan mencegah *over-fitting*.
- *Pemilihan Fitur Acak*
Pada setiap *node* dari *decision trees*, model akan mempertimbangkan semua fitur untuk menemukan pemisahan (*split*) terbaik. Hal ini membantu untuk mengurangi korelasi pohon-pohon dalam hutan, sehingga meningkatkan akurasi dan ketahanan model.

Karena sifatnya sebagai model *ensemble* yang kuat dan relatif mudah untuk disetel, RF telah banyak digunakan dalam berbagai aplikasi hidrologi, termasuk pemodelan suhu air danau (Heddam et al., 2020; Dyba et al., 2022).

Arsitektur Random Forest (Regresi)



Gambar 2.8 Struktur RF

2.9 Physics-Informed Machine Learning

Physics-Informed Machine Learning (PIML) adalah *framework* yang mengintegrasikan teori fisika ke dalam arsitektur model *machine learning* (Hao et al., 2023a). Teori fisika seringkali diekspresikan dalam bentuk Persamaan Diferensial Parsial (PDP). Hal ini bertujuan untuk menghasilkan model yang tidak hanya akurat secara prediktif tetapi juga konsisten dengan hukum-hukum fisika yang mendasarinya. Salah satu implementasi PIML yang paling populer adalah *Physics-Informed Neural Network* (PINN).

2.9.1 Physics-Informed Neural Network (PINN)

PINN adalah jaringan saraf yang dilatih untuk menyelesaikan PDP dengan memasukkan residu dari persamaan tersebut ke dalam *loss function* (Raissi et al., 2019). Misalkan sebuah sistem fisika dapat dideskripsikan oleh PDP umum berikut:

$$f(\mathbf{x}, t; u, \nabla u, \nabla^2 u, \dots; \lambda) = 0, \quad (\mathbf{x}, t) \in \Omega \times [0, T] \quad (2.9.18)$$

di mana $u(\mathbf{x}, t)$ adalah solusi yang dicari, f adalah operator diferensial, $(\mathbf{x}, t)$ adalah koordinat ruang dan waktu dalam domain $\Omega \times [0, T]$, dan λ adalah parameter fisika.

Dalam kerangka PINN, solusi $u(\mathbf{x}, t)$ diaproksimasi oleh sebuah jaringan saraf $u_{NN}(\mathbf{x}, t; \theta)$, dengan θ adalah parameter (bobot dan bias) dari jaringan tersebut. *Loss function* total (L) dari PINN kemudian dirumuskan sebagai jumlah terbobot dari beberapa komponen (Hao et al., 2023a):

$$L(\theta) = w_{data} L_{data} + w_{pde} L_{pde} \quad (2.9.19)$$

dimana:

- L_{data} adalah *data loss*, yang mengukur ketidaksesuaian antara prediksi jaringan u_{NN} dan data observasi yang tersedia.

- L_{pde} adalah *physics loss*, yang mengukur sejauh mana aproksimasi u_{NN} melanggar PDP
- L_{bc} adalah *boundary condition loss*, yang memastikan solusi memenuhi kondisi batas yang ditentukan.
- w_{data} , w_{pde} , dan w_{bc} adalah bobot yang menyeimbangkan kontribusi dari setiap komponen *loss*.

Dengan meminimalkan $L(\theta)$, jaringan saraf belajar untuk menemukan solusi yang secara simultan sesuai dengan data observasi dan mematuhi hukum fisika.

2.9.2 Mekanisme Perhitungan Turunan dengan Automatic Differentiation

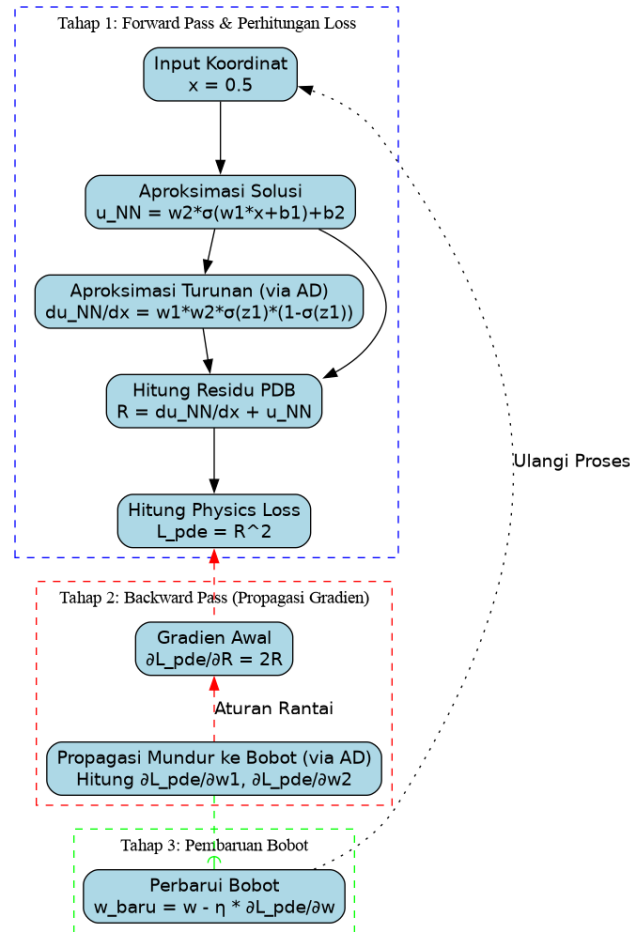
PINN bekerja dengan menggunakan Automatic Differentiation (AD) untuk menghitung turunan-turunan yang ada dalam residu PDP (L_{pde}). Seperti yang telah dijelaskan pada subbab sebelumnya, AD adalah teknologi yang memungkinkan perhitungan turunan dari fungsi yang kompleks terhadap *input* ($\mathbf{x}, t$) secara eksak.

Dengan AD, istilah-istilah turunan dalam PDP, seperti $\frac{\partial T}{\partial t}$ dan $\frac{\partial T}{\partial z}$ pada Persamaan 2.3.1, tidak perlu lagi diaproksimasi menggunakan metode numerik seperti beda hingga. Sebaliknya, kerangka kerja *deep learning* secara otomatis menghitung nilai eksak dari turunan-turunan ini pada setiap titik *collocation* selama proses pelatihan. Untuk turunan orde yang lebih tinggi seperti $\frac{\partial^2 T}{\partial z^2}$, dapat dihitung hanya dengan menerapkan AD secara berulang. Hal ini memungkinkan residu dari PDP untuk dimasukkan langsung ke dalam *loss function*. Sehingga dapat mendorong jaringan untuk tidak hanya cocok dengan data observasi tetapi juga mematuhi hukum fisika yang berlaku. Contoh alur AD dalam PINN disajikan dalam Gambar 2.9

2.9.3 Keunggulan PINN Dibandingkan Metode Numerik Tradisional

Model PINN memiliki beberapa keunggulan dibandingkan metode numerik tradisional seperti Metode Beda Hingga (*Finite Difference Method*, FDM) atau

Metode Elemen Hingga (*Finite Element Method*, FEM) (Hao et al., 2023a; Raissi et al., 2019).



Gambar 2.9 Contoh Automatic Differentiation (AD)

PINN mempelajari fungsi yang mendekati fungsi sebenarnya yang kontinu dari solusi. Setelah dilatih, model dapat dievaluasi pada setiap titik kontinu (t, z) di dalam domain tanpa perlu interpolasi. PINN bersifat *mesh-free* karena penggunaan AD tidak bergantung pada grid. PINN hanya memerlukan titik-titik *collocation* yang dapat diambil secara acak. Dalam kerangka PINN, parameter yang tidak diketahui seperti difusivitas τ dalam Persamaan 2.3.1 dapat didefinisikan sebagai variabel yang dapat dilatih. *Backpropagation* secara otomatis akan menghitung gradien dari *loss* terhadap parameter ini, dan optimizer akan mempelajarinya secara bersamaan dengan bobot jaringan. (Hao et al., 2023a; Raissi et al., 2019).

2.10 Aplikasi Deep Learning dalam Memprediksi Suhu Danau

Penggunaan model *Deep Learning* (DL) untuk memprediksi suhu danau telah berkembang pesat, dengan fokus utama pada kemampuannya untuk menangkap pola temporal yang kompleks dari data deret waktu. Berbagai arsitektur telah dieksplorasi, mulai dari model sekuensial seperti LSTM, model ekstraksi fitur seperti CNN, hingga pendekatan hibrida yang menggabungkan keduanya dengan pengetahuan fisika.

Model sekuensial, khususnya *Long Short-Term Memory* (LSTM), telah terbukti unggul dalam mempelajari dependensi jangka panjang pada data deret waktu suhu danau. Penelitian oleh Qiu et al. (2021) menunjukkan bahwa LSTM lebih akurat dibandingkan Random Forest (RF) dan *Back Propagation Neural Network* (BPNN). Aplikasinya bervariasi, mulai dari merekonstruksi data suhu historis selama 42 tahun di Danau-Danau Besar Amerika (Kayastha et al., 2023) hingga berfungsi sebagai model pengganti (*surrogate model*) yang ribuan kali lebih cepat dari model hidrodinamik konvensional (Wang et al., 2022). Arsitektur yang lainnya seperti *Entity-Aware LSTM* (EA-LSTM) bahkan mampu melakukan generalisasi prediksi ke danau-danau yang tidak termonitor dengan memasukkan fitur statis danau (Willard et al., 2022).

Convolutional Neural Network (CNN), meskipun awalnya untuk data citra, efektif dalam mengekstraksi fitur lokal dari data deret waktu melalui operasi konvolusi 1D. Zhu et al. (2024) berhasil menggunakan CNN untuk memprediksi profil suhu vertikal siklus harian, menangkap pola jangka pendek yang sering terlewatkan. Kekuatan CNN dalam ekstraksi fitur juga menjadikannya komponen ideal dalam arsitektur hibrida, seperti model CNN-LSTM yang dikembangkan oleh ?, yang terbukti lebih unggul daripada model tunggal.

Random Forest (RF) sering digunakan sebagai *benchmark* yang kuat karena kemampuannya menangani hubungan non-linear dengan penyetelan minimal (Qiu et al., 2021; Feigl et al., 2021). Studi perbandingan oleh Heddam et al. (2020) menemukan varian RF, yaitu *Extremely Randomized Trees* (ERT), sebagai model ML murni berkinerja terbaik. RF juga terbukti paling akurat untuk memperkirakan

suhu permukaan danau dari citra satelit (Dyba et al., 2022).

Untuk mengatasi keterbatasan model *black-box* yang terkadang tidak konsisten secara fisik, pendekatan *Physics-Informed Machine Learning* (PIML) atau *Process-Guided Deep Learning* (PGDL) menjadi tren utama. Pendekatan ini mengintegrasikan hukum fisika ke dalam model DL melalui dua strategi utama: *physics-informed loss* dan *pre-training*. Strategi pertama memberikan penalti pada model jika prediksinya melanggar hukum fisika, seperti konservasi energi (Read et al., 2019; Jia et al., 2020) atau stratifikasi densitas (Daw et al., 2021). Strategi kedua adalah melakukan pra-pelatihan model DL dengan data sintetis dari model berbasis proses (misalnya, GLM) untuk memberikan pemahaman dasar fisika sebelum disesuaikan dengan data observasi (Read et al., 2019; Jia et al., 2020). Model PIML/PGDL secara konsisten terbukti lebih akurat dan memiliki generalisasi yang lebih baik, terutama saat data terbatas (Read et al., 2019).

Dari tinjauan di atas, terlihat bahwa LSTM menjadi andalan untuk dependensi temporal, CNN untuk ekstraksi fitur, dan RF sebagai *benchmark* yang kuat. PIML muncul sebagai kerangka kerja yang menjanjikan untuk memastikan konsistensi fisika dan meningkatkan akurasi. Namun, sebagian besar penelitian masih berfokus pada prediksi suhu permukaan. Penelitian yang memodelkan profil suhu vertikal secara eksplisit, terutama dengan arsitektur hibrida seperti PINN-LSTM, merupakan area yang masih aktif dan menjanjikan, yang menjadi celah penelitian yang ingin diisi oleh studi ini. Tabel 2.1 merangkum studi-studi terkait yang relevan.

2.11 Metrik Penelitian

Untuk mengevaluasi kemampuan ML, terdapat berbagai metrik yang dirancang sebagai ukuran kuantitatif kinerja dari ML. Umumnya, ukuran kinerja ini spesifik untuk tujuan penelitian yang sedang dilakukan (Goodfellow et al., 2016). Metrik-metrik yang digunakan dalam penelitian ini adalah sebagai berikut.

- Root Mean Squared Error (RMSE): Metrik ini mengukur besarnya rata-rata kesalahan prediksi dalam satuan yang sama dengan variabel target (°C). Metrik

Tabel 2.1 Ringkasan Studi Terkait Aplikasi Machine Learning untuk Prediksi Suhu Air

Referensi	Input	Output	Metode	Domain
Read et al. (2019)	Data meteorologi (harian)	Profil suhu	PGDL (LSTM + GLM)	68 danau (Amerika Serikat)
Jia et al. (2020)	Data meteorologi (harian)	Profil suhu	PGRNN (LSTM + GLM)	Danau Mendota (Amerika Serikat)
Daw et al. (2021)	Data meteorologi, output GLM (harian)	Profil suhu	PGNN (MLP + GLM)	2 danau (Amerika Serikat)
Wang et al. (2022)	Data meteorologi, muka air (harian)	Suhu permukaan	LSTM (Surrogate)	2 waduk (Tiongkok)
Kayastha et al. (2023)	Data meteorologi (harian, lookback 5 hari)	Suhu permukaan (harian)	LSTM	Danau-Danau Besar Amerika Serikat
Willard et al. (2022)	Fitur statis danau, data meteorologi (harian)	Suhu permukaan (harian)	Entity-Aware LSTM	185.000 danau (Amerika Serikat)
Qiu et al. (2021)	Suhu udara, debit aliran (harian)	Suhu sungai (harian)	LSTM	9 sungai (global)
Zhu et al. (2024)	Suhu permukaan, suhu udara (per jam, lookback 24 jam)	Profil suhu (per jam)	CNN	Danau Kasumigaura (Jepang)
Barzegar et al. (2020)	Variabel kualitas air (15 menit, lookback 2 langkah)	DO, Chl-a (15 menit)	Hibrida CNN-LSTM	Danau Prespa (Yunani)
Dyba et al. (2022)	Citra Landsat 8 (pita spektral, indeks)	Suhu permukaan	Random Forest (RF)	38 danau (Polandia)
Heddham et al. (2020)	Suhu udara, komponen tanggal (harian)	Suhu permukaan (harian)	ERT, RF, MLP, dll.	25 danau (Polandia)
Feigl et al. (2021)	Data meteorologi, limpasan (harian, lookback 4 hari)	Suhu aliran (harian)	FNN, XGBoost, RNN	10 tambak ikan (Austria)
Di Nunno et al. (2023)	Suhu udara, LSWT tertunda (harian)	Suhu permukaan (prediksi 1–30 hari ke depan)	Stacked Ensemble (MLP, RF)	8 danau (Polandia)
Saber et al. (2020)	Fluks panas, angin, variabel tertunda (6 jam)	Profil suhu/DO (prediksi 240 hari ke depan)	ANN + Wavelet Transform	Danau Mead (Amerika Serikat)

ini digunakan oleh He and Yang (2025); Qiu et al. (2021); Heddham et al. (2020). RMSE memberikan bobot lebih pada kesalahan besar. Sehingga menonjolkan nilai prediksi yang berbeda jauh dari nilai aktual.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.11.20)$$

- Mean Absolute Error (MAE): Mengukur rata-rata kesalahan absolut antara nilai prediksi dan nilai sebenarnya. Metrik ini juga digunakan oleh He and Yang (2025); Qiu et al. (2021); Heddham et al. (2020). MAE kurang sensitif terhadap pencilan dibandingkan RMSE, karena perbedaan antara nilai prediksi dengan nilai aktual tidak dikuadratkan, sehingga *error* yang diberikan memiliki bobot yang sama.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.11.21)$$

- Nash-Sutcliffe Efficiency (NSE): Metrik ini membandingkan varians residual prediksi dengan varians data observasi, digunakan oleh He and Yang (2025); Qiu et al. (2021); Heddham et al. (2020). NSE yang mendekati angka 1 menunjukkan prediksi lebih baik, NSE yang bernilai 0 menunjukkan bahwa model sama akuratnya dengan rata-rata data, dan NSE kurang dari 0 menunjukkan bahwa rata-rata data adalah prediktor yang lebih baik.

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.11.22)$$

- Inkonsistensi Fisika (%): Metrik khusus yang dirancang untuk penelitian ini, mengukur persentase prediksi yang melanggar hukum fisika dasar stratifikasi termal. Metrik ini digunakan pada Daw et al. (2021); Jia et al. (2020); Read et al. (2019)

$$\text{Inkonsistensi Fisika} = \frac{\text{Jumlah prediksi tidak konsisten}}{\text{Total perbandingan antar lapisan}} \times 100\% \quad (2.11.23)$$

BAB III

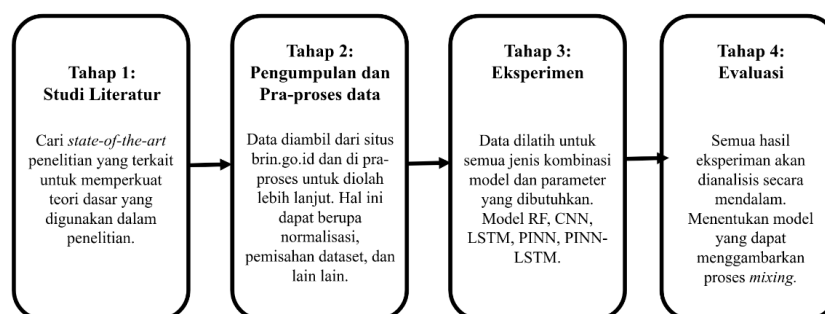
METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester ganjil tahun ajaran 2025/2026 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung yang beralamatkan di Jalan Prof.Dr.Ir.Soemantri Brojonegoro, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung secara *offline* dan di gedung Inderaja, KST Soekarno di Jl.Raya Bogor, Kecamatan Cibinong, Kabupaten Bogor, Jawa Barat secara *hybrid*. Fasilitas yang digunakan adalah BRIN HPC Mahameru CPU dan BRIN HPC Mahameru GPU yang dapat diakses secara *online*.

3.2 Metode Penelitian

Penelitian ini terdiri atas empat tahap. Adapun langkah-langkah yang dilakukan di dalam penelitian ini dapat dilihat pada Gambar 3.1:



Gambar 3.1 Metode Penelitian

Tahap studi literatur digunakan untuk mengetahui *state-of-the-art* dari penelitian yang terkait dan memperkuat landasan teori dari penelitian yang digunakan. Tahap

selanjutnya akan dikumpulkan data yang diperoleh untuk diproses terlebih dahulu sebelum digunakan dalam penelitian ini. Dalam tahap ini, pemrosesan data yang digunakan meliputi menyesuaikan tanggal, menggabungkan data, normalisasi, dan lain lain.

Pada tahap eksperimen, data yang sudah di pra-proses kemudian akan digunakan untuk melakukan eksperimen. tahap ini akan dilakukan dengan berbagai kombinasi eksperimen yang berbeda. Hasil dari tahap eksperimen akan dievaluasi dan dibandingkan dengan menggunakan metode evaluasi yang telah ditetapkan. Kombinasi yang memperoleh hasil yang terbaik akan dipilih untuk optimasi lebih lanjut.

Pada tahap terakhir, seluruh hasil eksperimen akan dianalisis secara mendalam. Tahapan ini juga akan membandingkan hasil eksperimen yang menggunakan prediksi model dengan hasil data yang sebenarnya dari metode observasi dan konvensional.

Komputasi dalam penelitian ini dilakukan menggunakan fasilitas MAHAMERU BRIN HPC GPU DGX, Badan Riset dan Inovasi Nasional (BRIN). Seluruh eksperimen implementasi *deep neural network* dilakukan pada mesin berbasis GPU dengan prosesor *Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz*, RAM 128GB, dan GPU *Nvidia Tesla V100 32GB*. Komputasi ini berbasis sistem operasi *Red Hat Enterprise Linux 8.8 (Ootpa)* dan dapat diakses melalui SSH seperti Gambar 3.2. Implementasi model pada penelitian ini adalah menggunakan bahasa pemrograman Python dan R dengan berbagai *library* seperti *Numpy* (Harris et al., 2020) *Tensorflow* (Abadi et al., 2015), *Sklearn* (Pedregosa et al., 2011), *Feather* (McKinney, 2016), dan *Pandas* (The pandas development team, 2025).

3.3 Studi Daerah Penelitian

Danau Batur merupakan danau kawah vulkanik terbesar di Bali dengan luas 16 km², kedalaman hingga 100 m, dan ketinggian 1.050 m. Peta danau batur tersaji dalam Gambar 3.3. Danau ini mendukung budidaya perikanan nila dan irigasi perairan (Wijaya et al., 2013).

```

=====
      Welcome to the MAHAMERU-AI 4th Gen. HPC System, nprayoga
      Computational services by BRIN
      https://hpc.brin.go.id | mailto:hpc@brin.go.id | wa:08119811568
=====

System Information:
Hostname      : mahoni01
Date & Time   : Tue Jan 27 23:02:50 WIB 2026
Uptime        : up 2 years, 12 weeks, 6 days, 21 hours, 9 minutes
Load Average: 5.64 5.81 5.94

Resource Usage:
Memory        : 148Gi / 251Gi used (59%)
Disk (home)   : 207T / 266T used (78%)

Cluster Utilization:
CPU Usage     : 6.5% (26/400 cores used)
Jobs          : 6 (running) | 2 (queued) | 8 (total)

Last Login:
nprayoga      pts/43    10.28.28.86    Tue Jan 27 23:02:50 +0700 2026

Your Active SLURM Jobs:
      JOBID PARTITION    NAME    USER ST      TIME  NODES NODELIST(REASON)

Tip: Remember to use 'srun' or 'sbatch' for compute jobs.

Researchers are requested to acknowledge this facility in
published works. For citation guidelines, please refer to:
https://hpc.brin.go.id/publications

Failure Notice: The DGX A100 is still under vendor maintenance,
please temporarily use the DGX A1, which has 8 Tesla V100 32 GB GPUs,
and Mahoni01, which has 2 Tesla V100 32 GB GPUs.

To ensure your job is submitted specifically to either A1 or Mahoni01,
use the following line in your SLURM script:
For A1      : #SBATCH --nodelist=a1
For Mahoni01: #SBATCH --nodelist=mahoni01
=====

```

Gambar 3.2 Laman SSH Mahameru BRIN HPC CPU

Namun, dalam 10 tahun terakhir status Danau Batur sudah berubah menjadi *hypereutrophic* akibat penumpukan nutrisi dari aktivitas pertanian dan antropogenik (Sunaryani et al., 2024; Nirasari et al., 2018). Kejadian ini menyebabkan pertumbuhan alga yang berlebih dan mengurangi kadar oksigen terlarut (Wetzel, 2001). Di sisi lain, *upwelling* merupakan pergerakan air dari dasar danau ke permukaan yang menyebabkan penurunan suhu dan pergerakan massa air kaya nutrisi ke permukaan, yang juga mengurangi kadar oksigen terlarut (DO) (Wetzel, 2001). Perubahan iklim, melalui peningkatan suhu udara dan perubahan kecepatan angin, juga menjadi kontributor besar terhadap perubahan suhu air danau (Wiradana et al., 2022).

Fenomena *upwelling* pada danau memiliki penyebab yang beragam, berdasarkan jenis danau itu sendiri. Danau Towuti, sebagai danau tropis lainnya di Indonesia. Danau ini merupakan danau air tawar terbesar kedua di Indonesia, dengan luas permukaan sekitar 560 km² dan kedalaman maksimum 203 m. Seperti Danau Batur, Danau Towuti juga merupakan danau terstratifikasi yang mengalami fenomena *upwelling* pada tahun 2008, yang disebabkan oleh perubahan suhu udara dan kecepatan angin (Pu et al., 2025). Sehingga karena kemiripan ini, dapat disimpulkan bahwa Danau Batur memiliki penyebab yang hampir sama seperti Danau Towuti.

Secara ekogeografis, Danau Batur merupakan kawah Gunung Batur yang pernah meletus dan memiliki kandungan logam berat. Logam berat dapat berasal dari pelapukan batuan yang mengandung kandungan seperti belerang. Sehingga ketika danau terjadi proses *upwelling*, zat-zat tersebut kemudian naik ke permukaan danau sehingga menghabiskan kadar oksiden dalam danau, menurunkan suhu air danau, dan meracuni danau tersebut dengan belerang. Logam berat merupakan salah satu polutan lingkungan yang bertanggung jawab atas toksisitas, ketahanan, dan sifat bioakumulasinya. Selain itu, akumulasi residu pestisida yang digunakan petani di sekitar Danau Batur untuk pemeliharaan sayuran juga dapat terbawa ke danau melalui aliran permukaan dan air hujan. Oleh karena itu, Menentukan kapan terjadinya proses *upwelling* merupakan suatu hal yang penting.

Prediksi *upwelling* pada Danau Batur dapat dijadikan sebagai bahan pertimbangan



Gambar 3.3 Peta danau Batur (Apple, 2025)

untuk para petani ikan lokal, pemerintah, dan otoritas terkait. Dengan model ini, petani ikan lokal dapat mengantisipasi kejadian *upwelling* sehingga dapat melakukan langkah-langkah pencegahan terhadap fenomena tersebut.

3.3.1 Dinamika Stratifikasi dan Pencampuran Danau Batur

Penelitian yang dilakukan Santoso et al. (2025a) yang menjelaskan tentang dinamika stratifikasi dan pencampuran Danau Batur menggunakan data suhu kolom air dan meteorologi frekuensi tinggi. Suhu permukaan danau berkisar antara 23.08-28.62°C, sementara suhu dasar (55 m) relatif stabil pada 23.11-23.64°C. Pengamatan ini mengkonfirmasi bahwa Danau Batur memiliki siklus polimiktik (*polymictic*) dengan empat periode musiman yang berbeda:

1. Stratifikasi awal (Oktober-Desember): Perkembangan stratifikasi termal.
2. Pencampuran periodik (Desember-Maret): Pencampuran yang dipicu oleh

kejadian angin kuat.

3. Stratifikasi kuat (Maret-Mei): Stabilitas termal maksimum.
4. Pencampuran sering (Mei-Oktober): Pencampuran nokturnal yang dipicu oleh pendinginan permukaan.

Pencampuran vertikal di Danau Batur dipicu oleh tiga faktor utama: kehilangan panas dari permukaan, kejadian angin kuat, dan pendinginan nokturnal (Santoso et al., 2025a). Pencampuran ini memiliki konsekuensi penting bagi ekosistem danau. Air dasar yang anoksik dan kaya sulfida dapat terangkat ke permukaan selama kejadian pencampuran. Studi tersebut memperkirakan bahwa sekitar 20% oksigen terlarut dalam kolom air dikonsumsi oleh oksidasi sulfida selama pencampuran, yang merupakan estimasi konservatif. Kondisi ini meningkatkan risiko kematian ikan massal, terutama bagi ikan yang dibudidayakan dalam Keramba Jaring Apung (KJA).

3.4 Data yang digunakan

Dalam penelitian ini, data yang digunakan merupakan data yang berformat CSV. CSV merupakan singkatan dari *Comma Seperated Values* adalah format yang sering digunakan untuk pertukaran data antar aplikasi yang berbeda. Format file ini telah menjadi standar *de facto* di seluruh industri (Shafranovich, 2005). Format ini biasanya dikeluarkan oleh alat-alat penelitian. Berikut merupakan struktur data CSV yang digambarkan pada Gambar 3.4.

Data prediktor yang digunakan didapatkan dari Badan Riset dan Inovasi Nasional (BRIN). Dataset ini dapat diambil dari Repositori Ilmiah Nasional (RIN) di BRIN (Santoso et al., 2025c). Dataset ini berupa data meteorologi yang diambil menggunakan HOBO U30 USB Weather Station. Dataset ini terdiri atas empat variabel, yaitu suhu udara (airT, °C, kelembapan relatif (RH, %), radiasi matahari (solar, Wm^{-2}), dan kecepatan angin (WindSpeed, ms^{-1}). Alat yang mengukur variabel ini dipasang di tepi Danau Batur. Semua data tersebut dihasilkan setiap satu

jam (Santoso et al., 2025c).

datetime,"wtr_0","wtr_2","wtr_4","wtr_6","wtr_8","wtr_10","wtr_15","wtr_20","wtr_30","wtr_55"
2022-10-08 00:00:00,24.363,24.098,24.098,24.05,24.05,24.05,23.785,23.617,23.521,23.424
2022-10-08 01:00:00,24.243,24.146,24.074,24.05,24.074,24.074,23.761,23.617,23.521,23.424
2022-10-08 02:00:00,24.122,24.219,24.171,24.05,24.002,24.002,23.761,23.617,23.497,23.424
2022-10-08 03:00:00,24.098,24.291,24.195,24.05,24.074,24.026,23.785,23.641,23.497,23.424
2022-10-08 04:00:00,24.05,24.388,24.243,24.122,24.074,24.002,23.809,23.665,23.545,23.424
2022-10-08 05:00:00,24.098,24.388,24.171,24.074,24.074,24.05,23.857,23.665,23.545,23.424
2022-10-08 06:00:00,24.243,24.508,24.219,24.074,24.074,24.002,23.833,23.665,23.545,23.424
2022-10-08 07:00:00,24.436,24.484,24.195,24.05,24.05,23.978,23.785,23.689,23.545,23.424
2022-10-08 08:00:00,24.919,24.484,24.195,24.05,24.074,24.002,23.785,23.737,23.545,23.424
2022-10-08 09:00:00,25.671,24.363,24.122,24.026,24.002,23.978,23.809,23.713,23.521,23.424
2022-10-08 10:00:00,25.866,24.267,24.098,24.026,24.002,23.978,23.809,23.713,23.521,23.424
2022-10-08 11:00:00,25.987,24.219,24.122,24.026,24.05,24.026,23.809,23.713,23.521,23.424
2022-10-08 12:00:00,25.574,24.412,24.098,23.978,23.905,23.905,23.737,23.689,23.569,23.424
2022-10-08 13:00:00,25.21,24.219,24.074,23.881,23.881,23.833,23.737,23.665,23.521,23.4
2022-10-08 14:00:00,25.016,24.195,24.05,23.881,23.881,23.905,23.737,23.641,23.521,23.4
2022-10-08 15:00:00,24.847,24.243,24.05,23.881,23.881,23.857,23.737,23.617,23.521,23.4
2022-10-08 16:00:00,24.508,24.291,24.05,23.954,23.905,23.881,23.689,23.593,23.497,23.4
2022-10-08 17:00:00,25.331,24.267,24.05,23.954,23.905,23.881,23.713,23.593,23.472,23.4
2022-10-08 18:00:00,25.623,24.195,24.026,23.905,23.905,23.833,23.713,23.593,23.472,23.4
2022-10-08 19:00:00,25.065,24.074,24.05,23.881,23.905,23.857,23.713,23.593,23.472,23.4
2022-10-08 20:00:00,24.484,24.002,24.002,23.881,23.905,23.857,23.713,23.617,23.472,23.4
2022-10-08 21:00:00,24.315,23.954,23.954,23.93,23.93,23.93,23.761,23.617,23.472,23.424
2022-10-08 22:00:00,24.171,24.122,24.122,24.074,24.026,24.026,23.833,23.737,23.569,23.424
2022-10-08 23:00:00,24.074,24.195,24.195,24.122,24.122,24.05,23.833,23.761,23.593,23.424
2022-10-09 00:00:00,23.978,24.146,24.171,24.122,24.098,24.026,23.857,23.785,23.617,23.424
2022-10-09 01:00:00,23.905,24.219,24.219,24.171,24.171,24.074,23.905,23.761,23.617,23.424
2022-10-09 02:00:00,24.074,24.219,24.195,24.122,24.074,23.954,23.857,23.713,23.641,23.424
2022-10-09 03:00:00,24.146,24.46,24.267,24.122,24.098,23.954,23.833,23.713,23.617,23.424
2022-10-09 04:00:00,24.146,24.557,24.412,24.122,24.122,24.122,23.93,23.809,23.641,23.424
2022-10-09 05:00:00,24.195,24.557,24.291,24.074,24.05,24.002,23.93,23.809,23.665,23.424
2022-10-09 06:00:00,24.243,24.508,24.267,24.05,24.026,24.05,23.978,23.881,23.737,23.424
2022-10-09 07:00:00,25.065,24.726,24.363,24.05,24.098,24.122,24.002,23.809,23.593,23.4

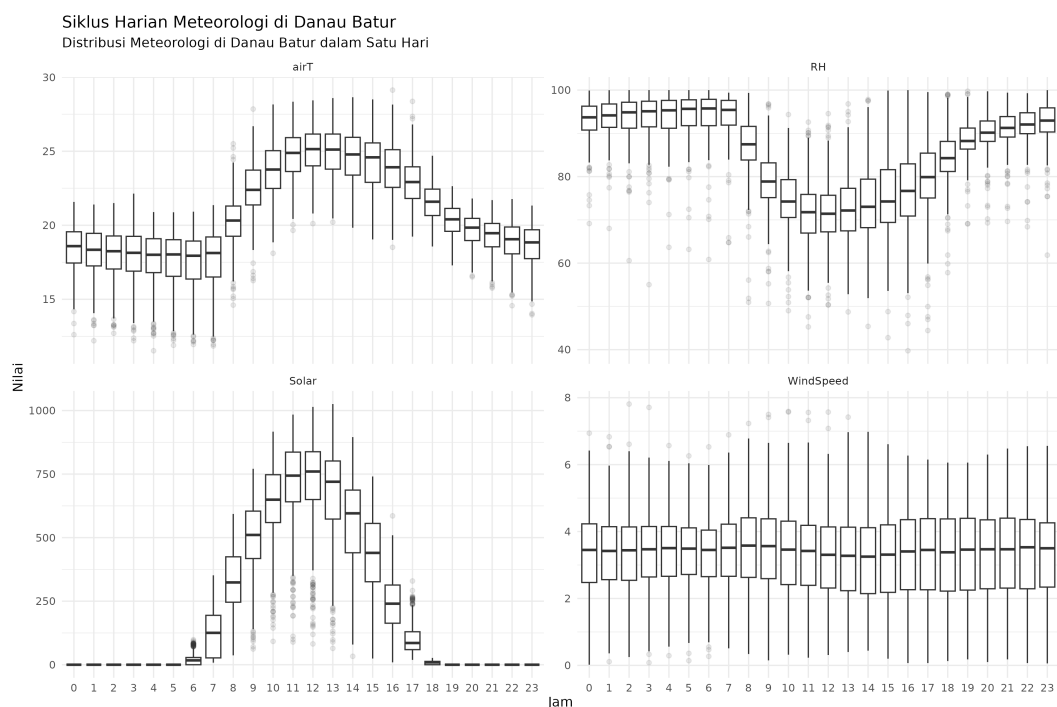
Gambar 3.4 Struktur data dari berkas CSV

Variabel suhu udara, kelembapan relatif, dan radiasi matahari memiliki siklus harian yang umumnya terjadi pada pertengahan hari, sedangkan kecepatan angin tidak memiliki siklus secara harian. Hal ini terlihat pada Gambar 3.5. Suhu udara dan radiasi matahari mengalami peningkatan, dan kelembapan relatif mengalami penurunan pada pertengahan hari. Kecepatan angin terlihat tidak memiliki siklus karena terdominasi oleh faktor lain seperti jet bertekanan rendah, sistem suhu tekanan rendah, dan aliran muson di daerah tropis pergunungan seperti Danau Batur (Nicolini

and Skabar, 2011).

Penelitian ini juga menggunakan variabel suhu air di berbagai titik kedalaman dari RIN sebagai *ground truth* target prediksi dan data prediktor sebagai observasi sebelumnya (Santoso et al., 2025b). Variabel ini diambil dengan serangkaian sensor suhu HOBO TidbiT V2 yang dipasang dari Oktober 2022 hingga Oktober 2023, sesuai dengan penelitian dari Santoso et al. (2025a). Perekaman ini dilakukan selama 1 jam sekali dengan akurasi $\pm 0,21$ °C dan resolusi 0,02 °C. Penelitian ini akan menggunakan tingkat kedalaman 0, 2, 4, 6, 8, 10, 15, 20, 30, dan 55 meter dari permukaan air.

Dataset ini direpresentasikan dalam sebuah tabel yang dimana setiap nilai dalam satu baris merepresentasikan sebuah nilai dalam satuan waktu. Baris pertama merepresentasikan judul atau variabel dari suatu kolom, sedangkan baris kedua dan selanjutnya merepresentasikan pengambilan variabel dalam satuan waktu.

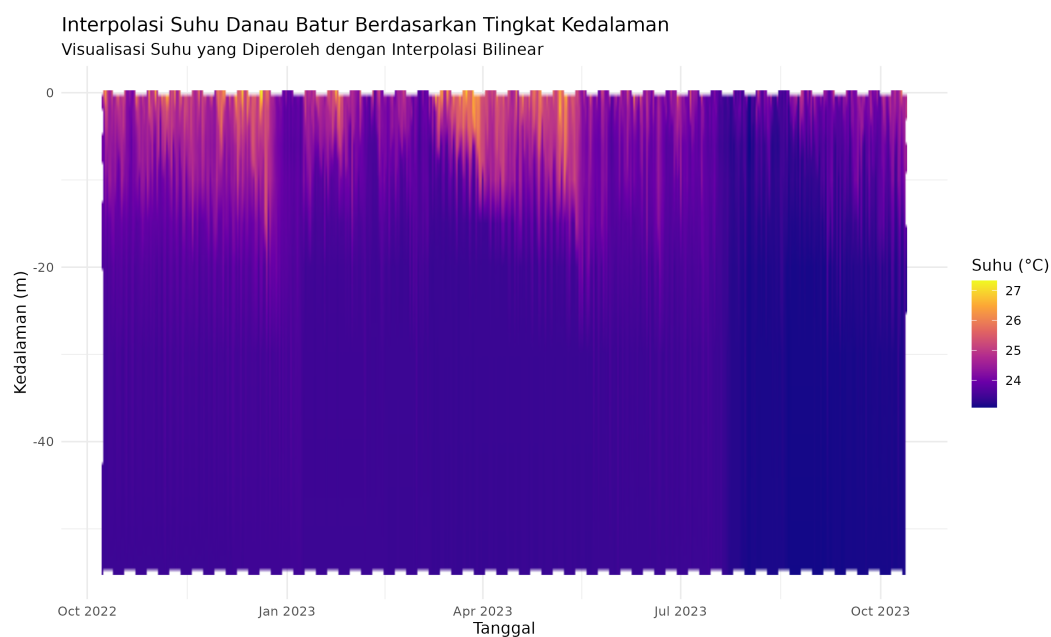


Gambar 3.5 Siklus harian meteorologi

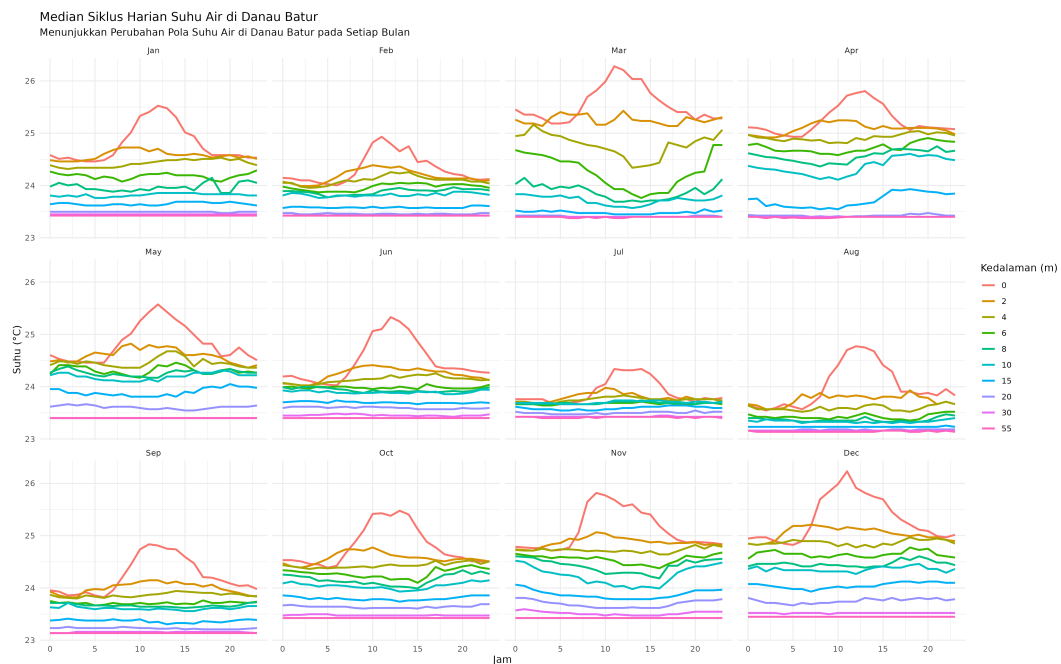
Variabel suhu permukaan dapat disajikan dalam bentuk *heatmap* pada Gambar 3.6 dengan rentang warna kuning-biru. Warna kuning menunjukkan tingkat kepanasan

yang lebih tinggi dibanding warna biru. Berdasarkan Gambar 3.6, terdapat rentang tanggal yang dimana suhu air di danau batur itu relatif homogen dari tingkat ketinggian teratas sampai terendah. Rentang tersebut terjadi pada bulan Januari 2023, Maret 2023, Juli 2023, dan September 2023. Hal ini sesuai dengan penelitian Pu et al. (2025) yang menunjukkan bahwa danau di daerah tropis memiliki siklus stratifikasi yang tidak teratur dan sering dipengaruhi oleh faktor meteorologi seperti suhu udara dan kecepatan angin.

Setiap lapisan ketinggian air di Danau batur memiliki siklus suhu yang berbeda. Hal ini terlihat pada Gambar 3.7. Dalam satu hari, selisih suhu di permukaan dan suhu di titik terdalam dapat mencapai 3.0°C pada pertengahan hari atau sampai dibawah 1.0°C pada malam hari. Secara umum, selisih terkecil dicapai dalam bulan juli dan selisih terbesar dicapai pada bulan maret.



Gambar 3.6 Data asli suhu Danau Batur diinterpolasi



Gambar 3.7 Siklus harian suhu danau batur dibagi berdasarkan bulan

3.4.1 Preprocessing Data

Terdapat tujuh tahap *preprocessing* yang perlu dilakukan agar data yang telah dikumpulkan dapat diproses oleh model. Gambar 3.8 menunjukkan seluruh alur tahap *preprocessing* data yang dilakukan dalam bahasa pemrograman R. Ketujuh tahap tersebut adalah sebagai berikut.

1. Tahap 1: Baca format input.

Pada tahap pertama, data-data seperti data meteorologi dan data suhu asli akan dibaca oleh komputer. Semua data ini dipresentasikan dalam sebuah tabel. Alat yang digunakan untuk membaca data ini adalah R.

2. Tahap 2: Penyesuaian tanggal.

Setelah semua data dibaca, akan dicari rentang tanggal yang sama untuk semua data yang diperoleh. Hal ini agar semua dataset memiliki rentang tanggal yang sama. Data yang diluar rentang tanggal yang sama akan di buang.

3. Tahap 3: Penggabungan data.

Data asli meteorologi dan profil suhu air yang keduanya dalam format CSV kemudian digabungkan menjadi satu tabel data berdasarkan penyesuaian tanggal di tahap sebelumnya. Data yang sudah digabungkan tersebut kemudian di urutkan secara kronologis. Detail masing-masing variabel dalam CSV dapat dilihat pada Tabel 3.1.

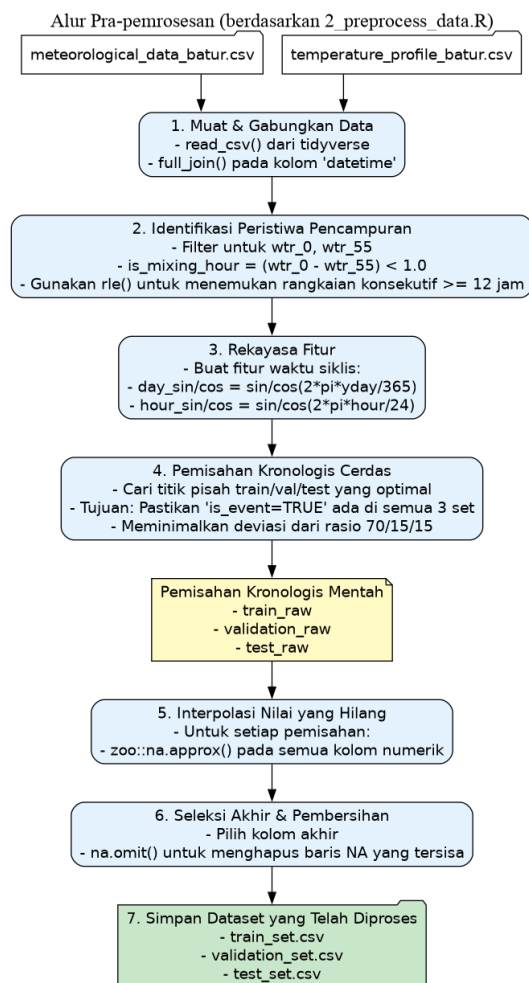
Variabel	<i>level</i>	Satuan	Singkatan
suhu udara	1	°C	airT
kelembapan relatif	1	%	RH
radiasi matahari	1	Wm^{-2}	solar
kecepatan angin	1	ms^{-1}	WindSpeed
suhu air	10	°C	wtr

Tabel 3.1 Detail variabel yang digunakan sebagai prediktor dalam penelitian ini

Pemilihan variabel meteorologi ini didasarkan pada dinamika stratifikasi dan pencampuran Danau Batur yang dijelaskan oleh Santoso et al. (2025a). Penelitian tersebut menunjukkan bahwa stabilitas stratifikasi dan pencampuran di Danau Batur sangat dipengaruhi oleh neraca panas permukaan dan gaya mekanis angin. Suhu udara dan radiasi matahari merupakan sumber panas utama yang menyebabkan stratifikasi termal pada siang hari melalui pemanasan permukaan (*surface heating*). Sebaliknya, kehilangan panas (*heat loss*) yang memicu pendinginan konvektif di malam hari sangat dipengaruhi oleh fluks panas laten dan sensibel, yang merupakan fungsi dari suhu udara, kelembapan relatif, dan kecepatan angin. Selain itu, kecepatan angin juga memberikan energi mekanis (*wind stress*) yang dapat memiringkan termoklin dan memicu pencampuran vertikal (*vertical mixing*), terutama pada periode angin kencang. Oleh karena itu, keempat variabel ini—suhu udara, kelembapan relatif, radiasi matahari, dan kecepatan angin—merupakan prediktor fisik yang esensial untuk memodelkan dinamika suhu danau.

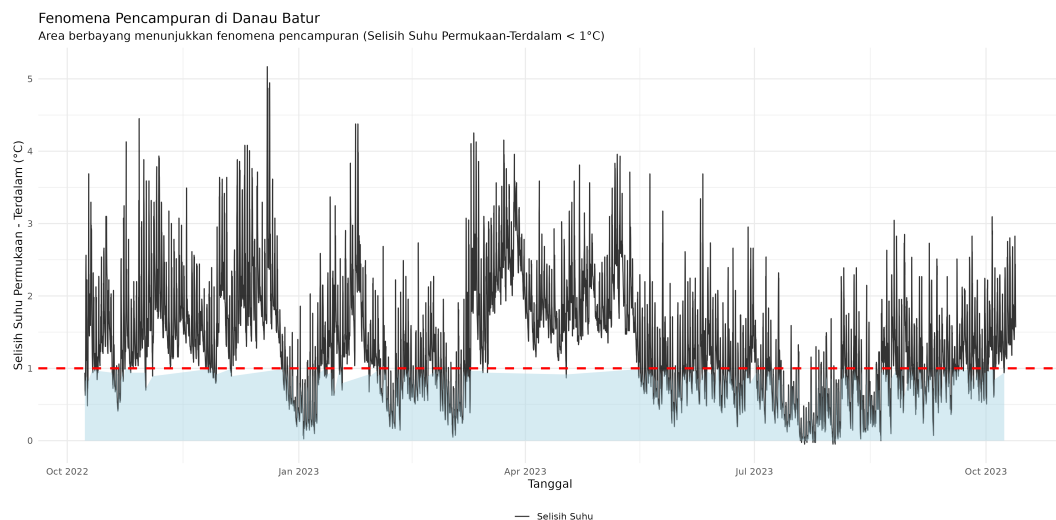
4. Tahap 4: Identifikasi fenomena *mixing* pada danau.

Peristiwa *mixing* didefinisikan sebagai periode berkelanjutan dengan periode minimal 12 jam dengan perbedaan suhu antara permukaan (0m) dan dasar danau



Gambar 3.8 Seluruh Tahap *Preprocessing* Data

(55m) kurang dari 1.0°C . Batas ini ditentukan atas pertimbangan dari beberapa artikel yang sudah dilakukan oleh penelitian lain (Engelhardt and Kirillin, 2014; Jung et al., 2023; Feldbauer et al., 2022; Schwefel et al., 2024). Ilustrasi untuk suhu yang termasuk dalam fenomena mixing tersaji pada Gambar 3.9



Gambar 3.9 Batas 1.0°C yang disebut sebagai fenomena mixing

5. Tahap 5: Pemilihan variabel.

Langkah selanjutnya adalah memilih variabel yang akan digunakan dalam penelitian ini. Sesuai dengan dijelaskan pada bagian sebelumnya, data prediktor yang digunakan adalah suhu udara, kelembapan relatif, radiasi matahari, kecepatan angin dan suhu air di Danau Batur dari periode sebelumnya.

Terdapat variabel tambahan yang digunakan yaitu *day of year* (DOY) dan *hour of day* (HOD). Menurut Heddham et al. (2020) dan Daw et al. (2021), variabel DOY memiliki pengaruh yang signifikan terhadap memprediksi suhu danau. Namun penelitian tersebut menggunakan dataset dengan observasi harian. Kekurangan lainnya adalah DOY merupakan fitur linear, sehingga akan terjadi diskontinuitas antara akhir tahun dengan awal tahun (Adams and Vamplew, 1998). Oleh karena itu, variabel DOY akan dibuat menjadi dua variabel dengan bantuan fungsi sinus dan cosinus yaitu DOY-cos dan DOY-sin. Penelitian Heddham et al. (2020) juga menunjukkan bahwa semakin tinggi resolusi feature tanggal, maka akan semakin

berpengaruh. Sehingga penelitian ini ditambahkan HOD yang dibagi menjadi dua yaitu HOD-cos dan HOD-sin.

Sebagai contoh, untuk waktu observasi tertentu, misal tanggal dan jam “2023-08-15 14:00:00”, transformasi fitur siklis dilakukan sebagai berikut. Tanggal 15 Agustus 2023 merupakan hari ke-227 dalam satu tahun, dan jam 14:00 merupakan jam ke-14 dalam satu hari. Transformasi ke fitur siklis menggunakan fungsi sinus dan cosinus:

$$\text{hour_sin} = \sin\left(\frac{2\pi \times 14}{24}\right) = \sin(3.665) = -0.500 \quad (3.4.1)$$

$$\text{hour_cos} = \cos\left(\frac{2\pi \times 14}{24}\right) = \cos(3.665) = -0.866 \quad (3.4.2)$$

$$\text{day_sin} = \sin\left(\frac{2\pi \times 227}{365}\right) = \sin(3.907) = -0.707 \quad (3.4.3)$$

$$\text{day_cos} = \cos\left(\frac{2\pi \times 227}{365}\right) = \cos(3.907) = -0.707 \quad (3.4.4)$$

Dengan transformasi ini, nilai-nilai yang dekat secara temporal juga dekat dalam ruang fitur. Contohnya pada jam 23:00 dan jam 00:00 memiliki nilai fitur siklus yang berdekatan, berbeda dengan representasi linear yang akan memberikan jarak numerik yang besar.

6. Tahap 6: Pembagian Data (*Smart Split*).

Langkah selanjutnya adalah membagi data menjadi set pelatihan (70%), validasi (15%), dan pengujian (15%). Berbeda dengan pembagian acak, penelitian ini menggunakan algoritma khusus. Algoritma ini secara iteratif mencari titik pemisahan data yang optimal untuk memastikan bahwa setiap set data (latih, validasi, dan uji) mempertahankan urutan kronologis dan mendapatkan representasi dari peristiwa *mixing* yang penting (Gusak et al., 2025). Hal ini penting untuk melatih model yang kuat dan dapat diuji pada fenomena tertentu.

7. Tahap 7: Normalisasi *feature matrix*.

feature matrix yang diperoleh akan dilakukan perubahan skala pada seluruh dengan menggunakan metode *Z-score normalization*. Hal ini perlu dilakukan

karena setiap variabel memiliki satuan yang berbeda-beda sebagaimana terlihat pada Tabel 3.1 dengan jumlah pencilaan yang besar (Amnuaylojaroen et al., 2024). Setelah tahap ini dilakukan, seluruh variabel akan memiliki nilai dengan rata-rata 0 dan standar deviasi 1 dan disimpan dalam bentuk *numpy array*. *feature matrix* yang belum dinormalisasi tetap disimpan untuk digunakan sebagai perhitungan pada bagian PB dan diubah menjadi *numpy array*.

3.5 Desain Eksperimen dan Tahap Latih Model

Tahap eksperimen dalam penelitian ini dirancang untuk mengevaluasi dan membandingkan kinerja dari berbagai model *machine learning* dalam memprediksi suhu Danau Batur. Eksperimen ini mencakup model-model standar serta model hibrida. Seluruh model dioptimalkan menggunakan *library* Optuna (Akiba et al., 2019) untuk mencari *hyperparameter* terbaik. *Hyperparameter* awal yang digunakan berdasarkan dari berbagai artikel yang menggunakan model seperti CNN, LSTM, RF, dan PINN. *Hyperparameter* tersebut kemudian diberikan variasi sebesar dua tingkatan di atas dan di bawah nilai literatur. Alur tahap latih model disajikan pada Gambar 3.10.

Pendekatan variasi dua tingkatan ini dipilih untuk menyeimbangkan antara eksplorasi dan eksploitasi dalam ruang pencarian hiperparameter. Variasi yang terlalu kecil (satu tingkat) dapat menyebabkan pencarian terjebak pada solusi lokal yang serupa dengan literatur, sementara variasi yang terlalu besar dapat memperlambat konvergensi dan menghasilkan konfigurasi yang tidak stabil. Dengan dua tingkat variasi, ruang pencarian cukup luas untuk menemukan konfigurasi yang optimal untuk karakteristik spesifik Danau Batur, namun tetap terarah berdasarkan pengetahuan sebelumnya dari literatur (Bergstra and Bengio, 2012).

Sebagai bagian penting dari desain eksperimen, penelitian ini secara sistematis mengevaluasi kinerja model di bawah berbagai konfigurasi. Untuk setiap arsitektur model, panjang sekuens *input* divariasikan dengan 24, 48, dan 72 jam. *Output* divariasikan dengan 6, 12, 18, dan 24 jam. Hal ini menghasilkan total 12 kombinasi

(3×4) untuk setiap arsitektur dasar, yang memungkinkan analisis komprehensif tentang sensitivitas model terhadap jumlah data historis dan durasi peramalan.

3.5.1 Model yang Diuji

Untuk memastikan perbandingan yang adil dan komprehensif, penelitian ini mengimplementasikan dan menguji lima strategi pemodelan yang berbeda. Pemilihan model didasarkan pada tinjauan literatur ekstensif terhadap studi prediksi suhu air sebelumnya:

1. **Random Forest (RF)**: Dipilih sebagai *benchmark* untuk model *machine learning* konvensional. RF terbukti tangguh terhadap *overfitting* dan mampu menangani hubungan non-linear yang kompleks, menjadikannya standar komparasi yang kuat dalam pemodelan hidrologi (Heddami et al., 2020; Dyba et al., 2022).
2. **Long Short-Term Memory (LSTM)**: Dipilih sebagai representasi *state-of-the-art* untuk *Deep Learning* pada data deret waktu. Kemampuannya mengingat informasi jangka panjang menjadikannya ideal untuk menangkap tren musiman dan inersia termal danau (Qiu et al., 2021; Kayastha et al., 2023).
3. **Convolutional Neural Network (CNN)**: Dipilih untuk mengevaluasi efektivitas ekstraksi fitur lokal. Studi terbaru menunjukkan CNN unggul dalam menangkap pola jangka pendek seperti siklus diel (harian) yang sering terlewatkan oleh model lain (Zhu et al., 2024).
4. **Physics-Informed Neural Network (PINN)**: Dimasukkan untuk mengatasi masalah inkonsistensi fisika pada model *data-driven*. Dengan mengintegrasikan persamaan difusi panas, PINN menjamin prediksi yang mematuhi hukum termodinamika (Read et al., 2019; Jia et al., 2020).
5. **PINN-LSTM**: Diajukan dalam penelitian ini sebagai pendekatan hibrida untuk

menyeimbangkan *trade-off* antara akurasi (kekuatan LSTM) dan konsistensi fisika (kekuatan PINN), mengisi celah yang belum sepenuhnya dieksplorasi oleh studi sebelumnya.

Setiap model dioptimalkan menggunakan Optuna untuk menemukan set hyperparameter terbaik untuk setiap kombinasi konfigurasi sekuens (T_{in} dan T_{out}). Empat model dasar yang digunakan sebagai perbandingan adalah sebagai berikut:

Tabel 3.2 Ringkasan Hyperparameter dari Literatur untuk Menjustifikasi Ruang Pencarian Optuna

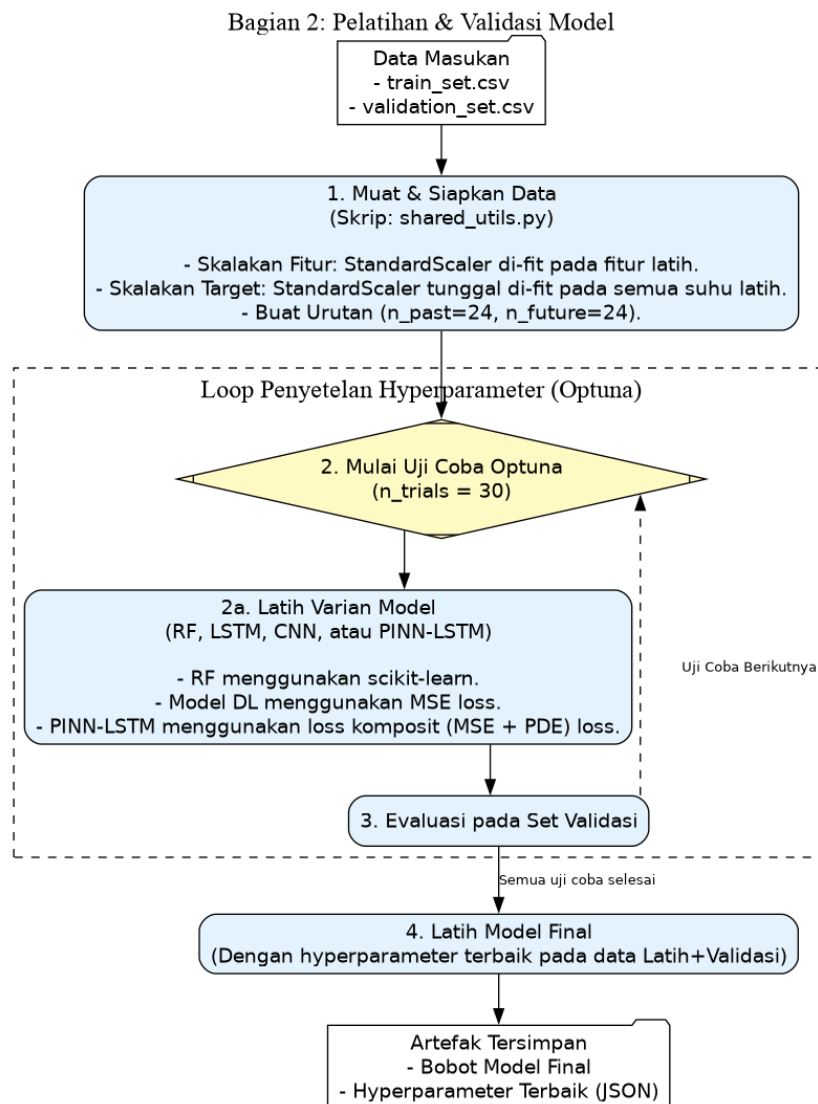
Model	Hyperparameter	Nilai	Sumber
PINN (MLP)	Arsitektur	4 <i>hidden layers</i>	He et al. (2025)
	<i>Tau Layers</i>	3	
	<i>Neuron per Layer</i>	50	
	Fungsi Aktivasi	<i>Softplus</i>	
	Optimizer	Adam (lr=0.001)	
LSTM	Hidden Units	21	Jia et al. (2021)
		30-50	Qiu et al. (2021)
	Optimizer	Adam (lr=0.005)	Jia et al. (2021)
CNN	Filters	64	Zhu et al. (2024)
	Kernel Size	2	
	Pooling Size	2	
	Arsitektur	Conv1D -> MaxPooling1D -> Dense	
	Optimizer	Adam (50 epochs, batch size 32)	
Random Forest	Estimators (Trees)	100-300	Qiu et al. (2021)
	Metode Seleksi Fitur	Gini impurity index	

- Random Forest (RF)

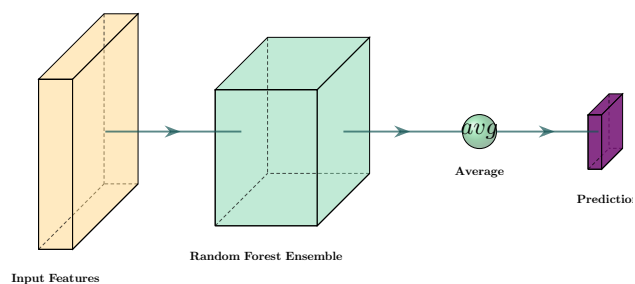
Model *ensemble* non-linear yang inputnya adalah data sekuens yang diratakan menjadi satu vektor fitur. Arsitektur RF disajikan pada Gambar 3.11.

- Long Short-Term Memory (LSTM)

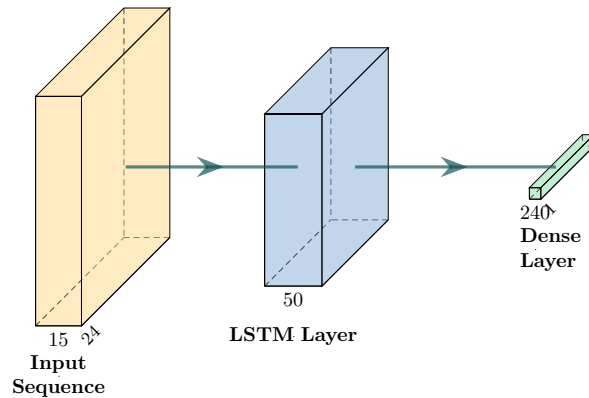
Jaringan saraf rekuren yang dirancang untuk data sekuensial. Inputnya adalah sekuens dengan panjang T_{in} dan 15 fitur. Arsitektur yang digunakan terdiri dari satu lapisan LSTM yang diikuti oleh lapisan *fully-connected* (Dense) untuk menghasilkan output prediksi berukuran $T_{out} \times 10$, seperti yang diilustrasikan pada Gambar 3.12.



Gambar 3.10 Seluruh Tahap latih model



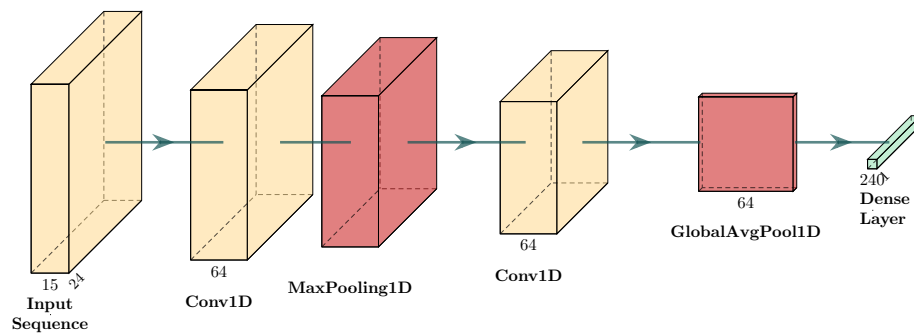
Gambar 3.11 Abstraksi Arsitektur Konseptual Model Random Forest.



Gambar 3.12 Arsitektur Model Long Short-Term Memory (LSTM).

- Convolutional Neural Network (CNN):

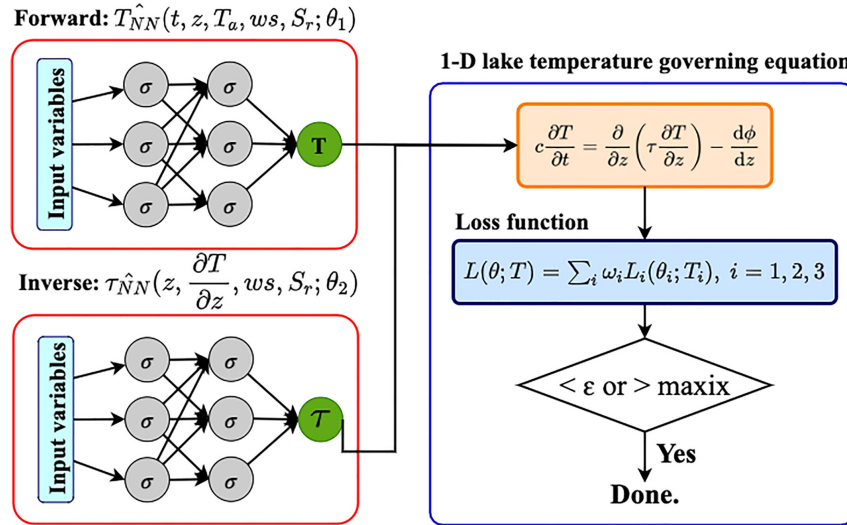
Model yang menggunakan lapisan konvolusi 1D untuk mengekstraksi fitur temporal dari sekuens input. Arsitektur yang diimplementasikan terdiri dari dua lapisan konvolusi yang diselingi dengan lapisan *max-pooling*, diikuti oleh lapisan *global average pooling* dan lapisan Dense untuk menghasilkan output berukuran $T_{out} \times 10$, yang ditunjukkan pada Gambar 3.13.



Gambar 3.13 Arsitektur Model Convolutional Neural Network (CNN).

- Physics Informed Neural Network (PINN):

Model yang menggunakan ANN sebagai dasar untuk memprediksi variabel. hasil prediksi tersebut akan dimasukkan dalam persamaan diferensial parsial untuk mengetahui ketepatan model dalam mengeluarkan hasil (He and Yang, 2025). Arsitektur PINN ditunjukkan pada Gambar 3.14.



Gambar 3.14 Arsitektur Model Physics Informed Neural Network (PINN) (He and Yang, 2025).

3.5.2 Model Hibrida PINN-LSTM

Model hibrida PINN-LSTM menjadi fokus utama dalam penelitian ini. Arsitektur ini menggabungkan dua model yang berbeda, yaitu LSTM dan PINN. Model ini terinspirasi oleh pendekatan dari Lawal et al. (2025), yang diilustrasikan pada Gambar 3.15, terdiri dari dua cabang utama yang bekerja secara sinergis:

- Cabang LSTM.

LSTM digunakan untuk peramalan deret waktu. Cabang ini menggunakan keunggulan LSTM dalam menangkap dependensi temporal untuk menghasilkan prediksi suhu di masa depan berdasarkan data historis. Kinerjanya diukur secara langsung melalui *data loss* terhadap data observasi.

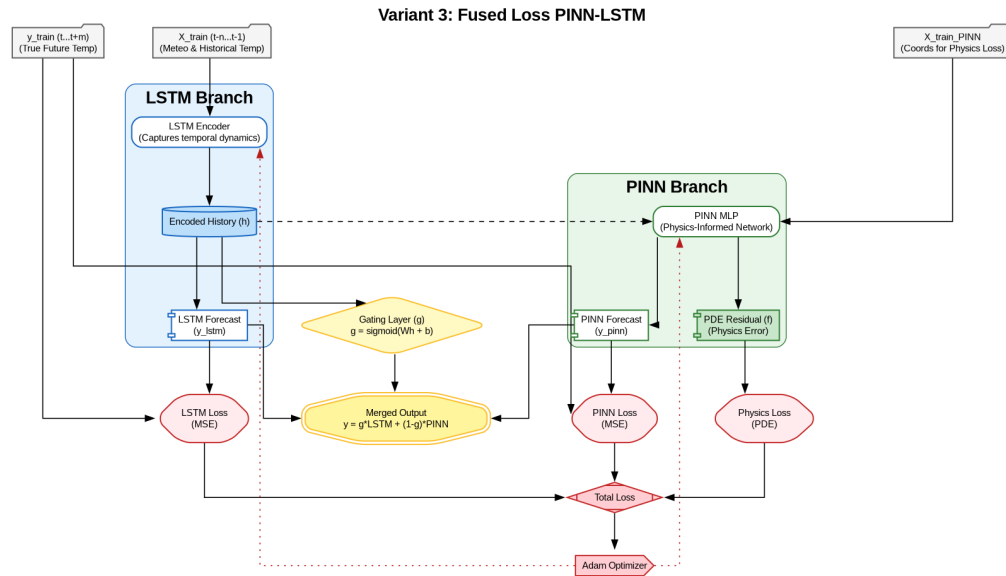
- Cabang PINN.

Berfungsi sebagai *regularizer* berbasis fisika. Cabang ini memiliki dua tujuan. Pertama, cabang ini juga dilatih untuk memprediksi suhu yang diukur dengan *data loss*-nya sendiri untuk mempelajari representasi data yang relevan. Kedua, cabang ini secara eksplisit mematuhi hukum fisika pada Persamaan 2.3.1 melalui *physics loss* yang dihitung pada titik-titik kolokasi (*collocation*).

Kemudian, terdapat Gated Fusion Mechanism. Lapisan *gating* (*g*) menentukan bobot

kontribusi dari cabang LSTM dan PINN untuk setiap prediksi. Dengan demikian, output akhir model (y_{merged}) merupakan gabungan dari prediksi berbasis data dan prediksi yang dipandu oleh fisika.

Loss function total dari model hibrida ini adalah jumlah dari ketiga komponen *loss* tersebut: *data loss* dari LSTM, serta *data loss* dan *physics loss* dari cabang PINN. Dengan melatih kedua cabang secara bersamaan, gradien dari *physics loss* dapat menginformasikan dan memperbaiki parameter di seluruh model, mendorong cabang LSTM utama untuk menghasilkan prediksi yang tidak hanya akurat tetapi juga lebih konsisten secara fisik.



Gambar 3.15 Arsitektur Model PINN-LSTM dengan Gated Fusion Mechanism.

3.5.3 Implementasi Sumber Radiasi Matahari

Dalam implementasi model PINN dan PINN-LSTM, sumber radiasi matahari ($\frac{d\phi}{dz}$) pada Persamaan 2.3.1 dihitung berdasarkan hukum Beer-Lambert. Turunan fluks radiasi terhadap kedalaman adalah:

$$\frac{d\phi}{dz} = -(1 - \beta)I_{sol}\eta e^{-\eta z} \quad (3.5.5)$$

Parameter β dan η ditetapkan sebagai konstanta tetap. Nilai tetapan yang digunakan dalam penelitian ini adalah $\beta = 0.4$ dan $\eta \approx 0.15 \text{ m}^{-1}$, yang konsisten dengan karakteristik kejernihan air Danau Batur. Hal ini dilakukan untuk memberikan kontribusi kepada *physics loss* dengan memastikan bahwa prediksi model mengikuti pola. Polanya berupa pemanasan maksimum yang terjadi di lapisan permukaan dan berkurang secara eksponensial dengan kedalaman. Dengan demikian, model tidak hanya mempelajari pola dari data, tetapi juga dibatasi oleh prinsip fisika transfer panas di dalam kolom air.

Untuk semua model, pencarian hiperparameter dilakukan menggunakan Optuna untuk menemukan kombinasi optimal yang meminimalkan *validation loss*. Ruang pencarian hiperparameter akan dibahas lebih lanjut pada bagian selanjutnya.

3.5.4 Optimasi Hiperparameter

Untuk setiap model, pencarian hiperparameter dilakukan secara otomatis menggunakan Optuna. Proses ini melibatkan pendefinisian ruang pencarian untuk setiap hiperparameter yang relevan (misalnya, *learning rate*, jumlah unit dalam lapisan, ukuran *kernel*, dll.) dan menjalankan 100 *trial* (percobaan) untuk menemukan kombinasi yang meminimalkan *validation loss*. Setiap *trial* akan dilakukan sebanyak 30 epoch dan ketika sudah menemukan hiperparameter yang terbaik, model akan melakukan pelatihan sebanyak 50 epoch.

Kombinasi model yang dibuat berdasarkan jumlah *input* kebelakang, dan jumlah *output* kedepan. Model akan memprediksi sejauh 1 jam, 6 jam, 12 jam, dan 24 jam kedepan, dengan menggunakan input 24 jam, 48 jam, dan 72 jam kebelakang. Sehingga satu arsitektur model memiliki $4 \times 3 = 12$ kombinasi. Total jumlah kombinasi model yang digunakan sebanyak 60 kombinasi. Tabel 3.3 menunjukkan hiperparameter yang akan dicari oleh Optuna.

Tabel 3.3 Ringkasan Hyperparameter yang digunakan sebagai Ruang Pencarian Optuna

Model	Hyperparameter	Nilai	Sumber
PINN (MLP)	<i>Hidden Layers</i>	2, 4, 6	He et al. (2025)
	<i>Tau Layers</i>	1, 2, 3, 4, 5	
	<i>Hidden Neuron per Layer (PINN)</i>	10, 20, 30, 40, 50, 60, 70, 80	
	<i>Tau Neuron per Layer (PINN)</i>	10, 20, 30, 40	
	Fungsi Aktivasi	<i>Softplus</i>	
	<i>Batch</i>	16, 32, 64	
LSTM	<i>Optimizer</i>	Adam (lr=[0.01-0.00001])	Jia et al. (2021)
	Hidden Units	16, 32, 48, 64, 80, 96, 112, 128	
	Fungsi Aktivasi	<i>Softplus</i>	
	<i>Batch</i>	16, 32, 64	
PINN-LSTM	<i>Optimizer</i>	Adam (lr=[0.01-0.00001])	Jia et al. (2021)
	<i>Hidden Layers (PINN)</i>	2, 4, 6	
	<i>Tau Layers (PINN)</i>	1, 2, 3, 4, 5	
	<i>Hidden Neuron per Layer (PINN)</i>	10, 20, 30, 40, 50, 60, 70, 80	
	<i>Tau Neuron per Layer (PINN)</i>	10, 20, 30, 40	
	Hidden Units (LSTM)	16, 32, 48, 64, 80, 96, 112, 128	
	Fungsi Aktivasi	<i>Softplus</i>	
	<i>Physics Loss Weight</i>	0.1-1.0	
CNN	<i>Batch</i>	16, 32, 64	Zhu et al. (2024)
	<i>Optimizer</i>	Adam (lr=[0.01-0.00001])	
	<i>Filters</i>	32, 64, 128	
	<i>Kernel Size</i>	2, 3, 5, 7	
	<i>Pooling Size</i>	2, 3, 5, 7	
	Arsitektur	Conv1D - MaxPooling1D - Dense	
Random Forest	<i>Batch</i>	16, 32, 64	Qiu et al. (2021)
	<i>Optimizer</i>	Adam (lr=[0.01-0.00001])	
	<i>Estimators (Trees)</i>	50, 100, 150, 200, 250, 300	
	Depth	10, 20, 30, 40, 50	
	<i>Minimum Sample for Leaf</i>	1-10	
	<i>Max Features Split</i>	0.2-1.0	
	Metode Seleksi Fitur	Gini impurity index	

3.5.5 Pelatihan dan Evaluasi

Setelah hiperparameter terbaik ditemukan melalui optimasi, setiap model dilatih kembali menggunakan gabungan data latih dan validasi. Kinerja akhir dari setiap model kemudian dievaluasi pada *test set* yang belum pernah dilihat sebelumnya. Proses ini memastikan perbandingan yang adil untuk semua model yang diuji. Artefak yang dihasilkan dari setiap proses latih model, termasuk bobot model final, hiperparameter terbaik, dan metrik kinerja pada test set, disimpan ke dalam berkas terpisah untuk dianalisis pada tahap selanjutnya.

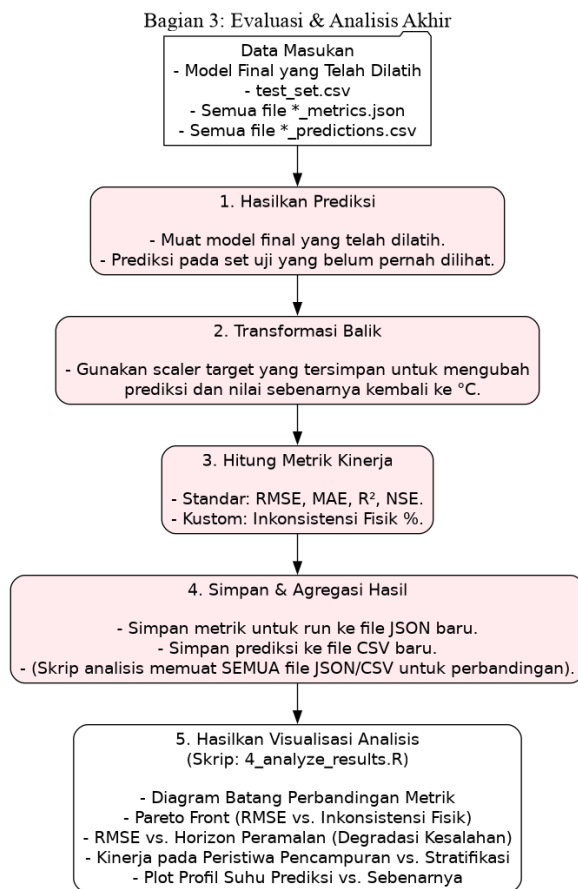
Selain metrik kualitas prediksi, efisiensi komputasi dari setiap model juga diukur secara eksplisit selama tahap ini. Dua metrik waktu dicatat menggunakan fungsi

`time.time()` pada Python untuk mengukur *wall-clock time*:

- **Waktu Pelatihan (Training Time):** Waktu total (dalam detik) yang dibutuhkan untuk melatih model final pada gabungan data latih dan validasi, mencakup semua epoch hingga kriteria pemberhentian terpenuhi.
- **Waktu Prediksi (Prediction Time):** Waktu total (dalam detik) yang dibutuhkan model yang telah dilatih untuk menghasilkan prediksi bagi keseluruhan *test set*. Metrik ini mengukur kecepatan inferensi model.

3.6 Metrik Evaluasi

Kinerja setiap model dievaluasi menggunakan serangkaian metrik standar dan metrik khusus fisika untuk memberikan penilaian yang komprehensif. Misalkan y_i adalah nilai suhu air observasi sebenarnya dan $\hat{y}_i$ adalah nilai yang diprediksi oleh model untuk sampel ke- i , dan n adalah jumlah total sampel. Seluruh tahap evaluasi disajikan pada Gambar 3.16. Metrik yang digunakan dalam penelitian ini adalah RMSE, MSE, NSE, dan Inkonsistensi Fisika.



Gambar 3.16 Seluruh Tahap Evaluasi Data

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Penelitian ini mengembangkan model prediksi suhu air Danau Batur menggunakan lima pendekatan *machine learning*: RF, LSTM, CNN, PINN, dan PINN-LSTM. Berdasarkan hasil eksperimen, dapat disimpulkan sebagai berikut:

1. **Model PINN-LSTM terbukti optimal** dengan RMSE 0.203°C , NSE 0.819, dan ketidakkonsistenan fisika hanya 3.42%. Model ini berhasil menggabungkan kekuatan LSTM dalam menangkap pola temporal dengan batasan fisika dari PINN.
2. **Terdapat *trade-off* signifikan** antara akurasi dan konsistensi fisika. Model data-driven (LSTM, CNN) mencapai RMSE rendah namun inkonsistensi tinggi (12-14%), sedangkan PINN konsisten secara fisika namun akurasi buruk (NSE negatif). Meskipun RF memiliki inkonsistensi rendah, model ini gagal menangkap dinamika pencampuran yang kritis. Hanya PINN-LSTM yang mampu menyeimbangkan akurasi dan fisika serta mendeteksi transisi ekstrim.
3. **Konfigurasi optimal**: 72 jam data historis ($n_p = 72$) dan prediksi 12 jam ke depan ($n_f = 12$). Penurunan akurasi konvergen pada jam ke-12, dan horizon ini memberikan waktu cukup untuk tindakan preventif.
4. **Periode stratifikasi** memiliki RMSE lebih tinggi karena gradien suhu besar antar kedalaman. Hanya PINN-LSTM yang mampu mereproduksi transisi stratifikasi-pencampuran dengan akurat dan konsisten.

5. **Model PINN murni tidak cocok** untuk Danau Batur karena persamaan difusi panas 1-D tidak cukup menjelaskan dinamika termal yang dipengaruhi angin, radiasi, dan aktivitas vulkanik.

5.2 Saran

Untuk pengembangan selanjutnya, disarankan:

1. **SOP Mitigasi Bencana KJA:** Mengembangkan prosedur standar operasional untuk petani Keramba Jaring Apung (KJA) ketika peringatan dini diterima (12 jam sebelum kejadian): (a) Pengurangan atau penghentian pemberian pakan untuk menurunkan kebutuhan oksigen, (b) Aktivasi aerator cadangan, (c) Jika memungkinkan, relokasi sementara ke zona aman yang lebih dangkal atau dekat pantai, dan (d) Pemanenan dini ikan yang sudah siap jual.
2. **Teknik *ensemble*:** Menggabungkan PINN-LSTM (unggul untuk transisi pencampuran) dengan LSTM (unggul untuk stratifikasi stabil) menggunakan *weighted averaging* adaptif.
3. **Validasi peristiwa ekstrim:** Kerja sama dengan Dinas Perikanan Kabupaten Bangli untuk memperoleh data peristiwa kematian ikan massal historis guna validasi sistem peringatan dini.
4. **Indeks peringatan dini:** Transformasi *output* prediksi menjadi kategori risiko (rendah-sangat tinggi) berdasarkan ambang batas perbedaan suhu permukaan-dasar.
5. **Optimasi arsitektur:** Eksplorasi mekanisme *attention*, *residual connections*, dan formulasi persamaan fisika yang mencakup efek angin dan radiasi matahari.
6. ***Transfer learning*:** Adaptasi model ke danau tropis lain seperti Danau Toba, Maninjau, atau Singkarak yang juga menghadapi masalah serupa.

DAFTAR PUSTAKA

Pengendalian pencemaran lingkungan danau batur bali dengan model decision support system (DSS), 2024. URL <https://brin.go.id/news/117709/pengendalian-pencemaran-lingkungan-danau-batur-bali-dengan-model-decision-support-system-dss>.

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL <https://www.tensorflow.org/>. Software available from tensorflow.org.

Anthony Adams and Peter Vamplew. Encoding and decoding cyclic data. *The South Pacific Journal of Natural Science*, 16, January 1998.

Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.

Solomon Oluwaseun Akinawo. Eutrophication: Causes, consequences, physical, chemical and biological techniques for mitigation strategies. *Environmental Challenges*, 12:100733, 2023. ISSN 2667-0100. doi: <https://doi.org/>

10.1016/j.envc.2023.100733. URL <https://www.sciencedirect.com/science/article/pii/S2667010023000574>.

Teerachai Amnuaylojaroen, Mariusz Ptak, and Mariusz Sojka. Assessment of the impact of meteorological variables on lake water temperature using the SHapley additive exPlanations method. *Water*, 16(22):3296, January 2024. ISSN 2073-4441. doi: 10.3390/w16223296.

Apple. Satellite view of Danau Batur, Bali. <https://maps.apple.com/>, 2025. Accessed: November 6, 2025.

Rahim Barzegar, Mohammad Taghi Aalami, and Jan Adamowski. Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stochastic Environmental Research and Risk Assessment*, 34(2):415–433, February 2020. ISSN 1436-3259. doi: 10.1007/s00477-020-01776-2.

Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.

James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13:281–305, 2012.

Bertram Boehrer and Martin Schultze. Stratification of lakes. *Reviews of Geophysics*, 46(3), 2008.

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Piyali Chowdhury, Naresh Kumar Goud Lakku, Susana Lincoln, Jaya Kumar Seelam, and Manasa Ranjan Behera. Climate change and coastal morphodynamics: Interactions on regional scales. *Science of The Total Environment*, 899:166432, 2023.

Arka Daw, Anuj Karpatne, William Watkins, Jordan Read, and Vipin Kumar. Physics-guided neural networks (PGNN): An application in lake temperature modeling, September 2021.

- Fabio Di Nunno, Senlin Zhu, Mariusz Ptak, Mariusz Sojka, and Francesco Granata. A stacked machine learning model for multi-step ahead prediction of lake surface water temperature. *Science of the Total Environment*, 890:164323, September 2023. ISSN 0048-9697. doi: 10.1016/j.scitotenv.2023.164323.
- Krzysztof Dyba, Sofia Ermida, Mariusz Ptak, Jan Piekarczyk, and Mariusz Sojka. Evaluation of methods for estimating lake surface water temperature using landsat 8. *Remote Sensing*, 14(15):3839, January 2022. ISSN 2072-4292. doi: 10.3390/rs14153839.
- Earthhow. How does eutrophication work? Causes, process and examples, August 2017.
- Jeffrey L Elman. Finding structure in time. *Cognitive science*, 14(2):179–211, 1990.
- Christof Engelhardt and Georgiy Kirillin. Criteria for the onset and breakup of summer lake stratification based on routine temperature measurements. *Fundamental and Applied Limnology / Archiv Für Hydrobiologie*, 184, June 2014. doi: 10.1127/1863-9135/2014/0582.
- J. Fajar. Fenomena mati masal ikan mujair di danau batur, August 2021. URL <https://mongabay.co.id/2021/08/03/fenomena-mati-masal-ikan-mujair-di-danau-batur/>.
- Moritz Feigl, Katharina Lebedzinski, Mathew Herrnegger, and Karsten Schulz. Machine-learning methods for stream water temperature prediction. *Hydrology and Earth System Sciences*, 25(5):2951–2977, May 2021. ISSN 1027-5606. doi: 10.5194/hess-25-2951-2021.
- J. Feldbauer, J. P. Mesman, T. K. Andersen, and R. Ladwig. Learning from a large-scale calibration effort of multiple lake temperature models. *Hydrology and Earth System Sciences*, 29:1183–1199, 2025. doi: 10.5194/hess-29-1183-2025. URL <https://doi.org/10.5194/hess-29-1183-2025>.
- Johannes Feldbauer, Robert Ladwig, Jorrit Mesman, Tadhg Moore, Hilke Zündorf, Thomas Berendonk, and Thomas Petzoldt. Ensemble of models shows coherent

response of a reservoir's stratification and ice cover to climate warming. *Aquatic Sciences*, 84, October 2022. doi: 10.1007/s00027-022-00883-2.

Felix A. Gers, Jürgen Schmidhuber, and Fred Cummins. Learning to forget: Continual prediction with lstm. *Neural Computation*, 12(10):2451–2471, 10 2000. ISSN 0899-7667. doi: 10.1162/089976600300015015. URL <https://doi.org/10.1162/089976600300015015>.

Malgorzata Golub, Wim Thiery, Rafael Marcé, Don Pierson, Inne Vanderkelen, Daniel Mercado-Bettin, R. Iestyn Woolway, Luke Grant, Eleanor Jennings, Benjamin M. Kraemer, Jacob Schewe, Fang Zhao, Katja Frieler, Matthias Mengel, Vasili Y. Bogomolov, Damien Bouffard, Marianne Côté, Raoul Couture, Andrey V. Debolskiy, and Galina Zdorovennova. A framework for ensemble modelling of climate change impacts on lakes worldwide: the isimip lake sector. *Geoscientific Model Development*, 15(11):4597–4623, 2022. doi: 10.5194/gmd-15-4597-2022. URL <https://doi.org/10.5194/gmd-15-4597-2022>.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

M. Guo, J. M. Melack, W. Zhou, P. M. Barbosa, J. H. F. Amaral, and Q. Zhuang. Linking biogeochemical and hydrodynamic processes to model methane fluxes in shallow, tropical floodplain lakes. *Journal of Advances in Modeling Earth Systems*, 15(12), 2023. doi: 10.1029/2022ms003385. URL <https://doi.org/10.1029/2022ms003385>.

Mingyang Guo, Qianlai Zhuang, Huaxia Yao, Malgorzata Golub, L. Ruby Leung, Don Pierson, and Zeli Tan. Validation and sensitivity analysis of a 1-d lake model across global lakes. *Journal of Geophysical Research: Atmospheres*, 126 (4), 2020. doi: 10.1029/2020jd033417. URL <https://doi.org/10.1029/2020jd033417>.

Danil Gusak, Anna Volodkevich, Anton Klenitskiy, Alexey Vasilev, and Evgeny Frolov. Time to split: Exploring data splitting strategies for offline evaluation of

sequential recommenders. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems*, pages 874–883, September 2025. doi: 10.1145/3705328.3748164.

Zhongkai Hao, Songming Liu, Yichi Zhang, Chengyang Ying, Yao Feng, Hang Su, and Jun Zhu. Physics-informed machine learning: A survey on problems, methods and applications, March 2023a.

Zihan Hao, Weide Li, Jinran Wu, Shaotong Zhang, Shujuan Hu, Zihan Hao, Weide Li, Jinran Wu, Shaotong Zhang, and Shujuan Hu. A novel deep learning model for mining nonlinear dynamics in lake surface water temperature prediction. *Remote Sensing*, 15(4), February 2023b. ISSN 2072-4292. doi: 10.3390/rs15040900.

Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2. URL <https://doi.org/10.1038/s41586-020-2649-2>.

Yuan He and Xiaofan Yang. A physics-informed deep learning framework for estimating thermal stratification in a large deep reservoir. *Water Resources Research*, 61(7):e2025WR040592, 2025. ISSN 1944-7973. doi: 10.1029/2025WR040592.

Salim Heddami, Mariusz Ptak, and Senlin Zhu. Modelling of daily lake surface water temperature from air temperature: Extremely randomized trees (ERT) versus Air2Water, MARS, M5Tree, RF and MLPNN. *Journal of Hydrology*, 588:125130, September 2020. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2020.125130.

M. R. Hipsey, L. C. Bruce, C. Boon, B. Busch, C. C. Carey, D. P. Hamilton, P. C.

- Hanson, J. S. Read, E. de Sousa, M. Weber, and L. A. Winslow. A general lake model (glm 3.0) for linking with high-frequency sensor data from the global lake ecological observatory network (gleon). *Geoscientific Model Development*, 12: 473–523, 2019. doi: 10.5194/gmd-12-473-2019. URL <https://doi.org/10.5194/gmd-12-473-2019>.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.
- Jörg Imberger and George T. Csanady. *Physical Limnology*. Elsevier, Amsterdam, 2015.
- Xiaowei Jia, Jared Willard, Anuj Karpatne, Jordan S. Read, Jacob A. Zwart, Michael Steinbach, and Vipin Kumar. Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles. 2:1–26, January 2020. doi: 10.1145/3447814.
- Eunsong Jung, Gea-Jae Joo, Hyo Gyeom Kim, Dong-Kyun Kim, and Hyun Woo Kim. Effects of seasonal and diel variations in thermal stratification on phytoplankton in a regulated river. *Sustainability*, 15:16330, November 2023. doi: 10.3390/su152316330.
- Miraj B. Kayastha, Tao Liu, Daniel Titze, Timothy C. Havens, Chenfu Huang, and Pengfei Xue. Reconstructing 42 years (1979–2020) of great lakes surface temperature through a deep learning approach. *Remote Sensing*, 15(17):4253, January 2023. ISSN 2072-4292. doi: 10.3390/rs15174253.
- Zaharaddeen Karami Lawal, Hayati Yassin, Daphne Teck Ching Lai, and Azam Che Idris. Physics-informed neural network (PINN) evolution and beyond: A systematic literature review and bibliometric analysis. *Big Data and Cognitive Computing*, 6(4):140–163, November 2022. ISSN 2504-2289. doi: 10.3390/bdcc6040140.
- Zaharaddeen Karami Lawal, Hayati Yassin, Daphne Teck Ching Lai, and Azam Che Idris. Modeling the complex spatio-temporal dynamics of ocean wave parameters:

- A hybrid PINN-LSTM approach for accurate wave forecasting. *Measurement*, 252:117383, 2025. doi: 10.1016/j.measurement.2025.117383.
- Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11): 2278–2324, 1998.
- Randall J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. Society for Industrial and Applied Mathematics, 2007.
- Yi Li, Dmitry Beletsky, Jia Wang, Jay Austin, John Kessler, Ayumi Fujisaki-Manome, and Peng Bai. Modeling a large coastal upwelling event in lake superior. *Journal of Geophysical Research: Oceans*, 126(3):e2020JC016512, 2021. doi: 10.1029/2020JC016512. URL <https://doi.org/10.1029/2020JC016512>.
- Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
- Wes McKinney. Feather: fast, interoperable binary data frame storage for Python, R, and more powered by Apache Arrow. <https://github.com/wesm/feather>, 2016.
- Daniel Mercado-Bettín, Francois Clayer, Muhammed Shikhani, Tadhg N. Moore, María Dolores Frías, Leah Jackson-Blake, James Sample, Maialen Iturbide, Sixto Herrera, Andrew S. French, Magnus Dahler Norling, Karsten Rinke, and Rafael Marcé. Forecasting water temperature in lakes and reservoirs using seasonal climate prediction. *Water Research*, 201:117286, August 2021. ISSN 0043-1354. doi: 10.1016/j.watres.2021.117286.
- Merriam-Webster. Upwelling, 2025. URL <https://www.merriam-webster.com/dictionary/upwelling>.
- Beat Müller, Lee D. Bryant, Andreas Matzinger, and Alfred Wüest. Hypolimnetic oxygen depletion in eutrophic lakes. *Environmental Science & Technology*, 46(18):9964–9971, September 2012. ISSN 0013-936X. doi: 10.1021/es301422r.

- Matilde Nicolini and Yanina García Skabar. Diurnal cycle in convergence patterns in the boundary layer east of the Andes and convection. *Atmospheric Research*, 100 (4):377–390, June 2011. ISSN 0169-8095. doi: 10.1016/j.atmosres.2010.09.019.
- K. G. Nirasari, I. W. Arya, and S. A. M. P. Suryani. Studi struktur komunitas fitoplankton di danau batur, kecamatan kintamani, kabupaten bangli, provinsi bali. *Gema Agro*, 23(1):104–107, 2018. doi: 10.22225/ga.23.1.664.104-107. URL <https://doi.org/10.22225/ga.23.1.664.104-107>.
- Ted Ozersky, Amanda Poste, Milla Rautio, and Eva Leu. Impacts of changing winters on lake ecosystems will increase with latitude. *Ecology Letters*, 28(8): e70200, 2025.
- Florian Pappenberger, Hannah L. Cloke, David J. Parker, Fredrik Wetterhall, David S. Richardson, and Jutta Thielen. The monetary benefit of early flood warnings in europe. *Environmental Science & Policy*, 51:278–291, 2015. doi: 10.1016/j.envsci.2015.04.016.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- Posmerdeka. 70 ton ikan di danau batur mati, tebar bau busuk ke permukiman warga, July 2025. URL <https://posmerdeka.com/70-ton-ikan-di-danau-batur-mati-tebar-bau-busuk-ke-permukiman-warga/>.
- T. Pu, G. D. Haffner, S. A. Crowe, and S. Katsev. Stratification stability of tropical lakes and their sensitivity to climate. *Limnology and Oceanography*, 2025. doi: 10.1002/lno.70055. URL <https://doi.org/10.1002/lno.70055>.
- Rujian Qiu, Yuankun Wang, Bruce Rhoads, Dong Wang, Wenjie Qiu, Yuwei Tao, and Jichun Wu. River water temperature forecasting using a deep learning method.

- Journal of Hydrology*, 595:126016, April 2021. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2021.126016.
- Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- Jordan S. Read, Xiaowei Jia, Jared Willard, Alison P. Appling, Jacob A. Zwart, Samantha K. Oliver, Anuj Karpatne, Gretchen J. A. Hansen, Paul C. Hanson, William Watkins, Michael Steinbach, and Vipin Kumar. Process-guided deep learning predictions of lake water temperature. *Water Resources Research*, 55(11): 9173–9190, 2019. doi: 10.1029/2019wr024922. URL <https://doi.org/10.1029/2019wr024922>.
- D. C. Roberts, G. C. Egan, A. L. Forrest, J. L. Largier, F. A. Bombardelli, B. E. Laval, S. G. Monismith, and G. Schladow. The setup and relaxation of spring upwelling in a deep, rotationally influenced lake. *Limnology and Oceanography*, 66(4):1168–1189, 2021. doi: 10.1002/lno.11673. URL <https://doi.org/10.1002/lno.11673>.
- Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
- David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986. doi: 10.1038/323533a0. URL <https://doi.org/10.1038/323533a0>.
- Ali Saber, David E. James, and Donald F. Hayes. Long-term forecast of water temperature and dissolved oxygen profiles in deep lakes using artificial neural networks conjugated with wavelet transform. *Limnology and Oceanography*, 65(6):1297–1317, 2020. ISSN 1939-5590. doi: 10.1002/lno.11390.
- Arianto B. Santoso, David P. Hamilton, Sulung Nomosatryo, Astried Sunaryani, Hadiid A. Rustini, Aldiano Rahmadya, Fajar Setiawan, Endra Triwisesa, Meti

Yulianti, Furqon Hensan Muttaqien, and Alifia Imananda. Stratification and mixing dynamics in tropical polymictic Lake Batur, Indonesia. *Inland Waters*, December 2025a. ISSN 2044-2041.

Arianto B. Santoso, Sulung Nomosatryo, Astried Sunaryani, Hadiid A. Rustini, Aldiano Rahmadya, Fajar Setiawan, Endra Triwisesa, Meti Yulianti, Furqon H. Muttaqien, Alifia Imananda, Ni Luh Kartini, and Farham Harvianto. Hourly Temperature Profile Dataset for Lake Batur, 2025b. URL <https://hdl.handle.net/20.500.12690/RIN/HN7JZP>.

Arianto B. Santoso, Sulung Nomosatryo, Astried Sunaryani, Hadiid A. Rustini, Aldiano Rahmadya, Fajar Setiawan, Endra Triwisesa, Meti Yulianti, Furqon H. Muttaqien, Alifia Imananda, Ni Luh Kartini, and Farham Harvianto. Hourly Meteorological Dataset for Lake Batur, 2025c. URL <https://hdl.handle.net/20.500.12690/RIN/SCO8RS>.

Robert Schwefel, Lipa Nkwale, Sylvia Jordan, Karsten Rinke, and Michael Hupfer. Temperatures and hypolimnetic oxygen in german lakes: Observations, future trends and adaptation potential. *Ambio*, 54, July 2024. doi: 10.1007/s13280-024-02046-z.

Yakov Shafranovich. Common format and mime type for comma-separated values (csv) files. 2005.

Tetsuya Shintani, Alberto de la Fuente, Yarko Niño, and Jörg Imberger. Generalizations of the wedderburn number: Parameterizing upwelling in stratified lakes. *Limnology and Oceanography*, 55(3):1377–1387, 2010. doi: 10.4319/lo.2010.55.3.1377. URL <https://doi.org/10.4319/lo.2010.55.3.1377>.

PC Sonarghare, SC Masram, UR Sonparote, KP Khaparde, and SK Kharkate. Causes and effects of eutrophication on aquatic life (a review). *International Journal for Environmental Rehabilitation and Conservation*, XI, pages 213–218, 2020.

- Peiyang Su, Jiayong Zhang, and Liwen Guo. A PINN-LSTM model for real-time gas concentration prediction in mine conveyor belt fires. *Fire*, 8:450, November 2025. doi: 10.3390/fire8110450.
- A Sunaryani, A B Santoso, F H Muttaqien, F Harvianto, H A Rustini, N L Kartini, D Nilawati, and E Triwisesa. A web-based numerical model for sustainable management of lake batur, bali: Preliminary results. *IOP Conference Series: Earth and Environmental Science*, 1201(1):12049, June 2023. ISSN 1755-1315. doi: 10.1088/1755-1315/1201/1/012049.
- A. Sunaryani, A. B. Santoso, P. Soewondo, N. Suharyanto, A. Imananda, and I. F. Sani. Eutrophication in lake batur: Current status and management strategies. *E3S Web of Conferences*, 485:03013, 2024. doi: 10.1051/e3sconf/202448503013. URL <https://doi.org/10.1051/e3sconf/202448503013>.
- D. F. Swinehart. The beer-lambert law. *Journal of Chemical Education*, 39(7): 333, 1962. doi: 10.1021/ed039p333. URL <https://doi.org/10.1021/ed039p333>.
- The pandas development team. pandas-dev/pandas: Pandas (v2.3.2), 2025. URL <https://doi.org/10.5281/zenodo.16918803>.
- United Nations Office for Disaster Risk Reduction. Global status of multi-hazard early warning systems: Target g, 2023. URL <https://www.undrr.org/reports/global-status-multi-hazard-early-warning-systems-2023>.
- Longfan Wang, Bo Xu, Chi Zhang, Guangtao Fu, Xiaoxian Chen, Yi Zheng, and Jingjie Zhang. Surface water temperature prediction in large-deep reservoirs using a long short-term memory model. *Ecological Indicators*, 134:108491, January 2022. ISSN 1470-160X. doi: 10.1016/j.ecolind.2021.108491.
- Xiwen Wang, Kun Shi, Yunlin Zhang, Boqiang Qin, Yibo Zhang, Weijia Wang, R. Iestyn Woolway, Shilong Piao, and Erik Jeppesen. Climate change drives rapid

warming and increasing heatwaves of lakes. *Science Bulletin*, 68(15):1574–1584, 2023.

Robert G. Wetzel. *Limnology: Lake and River Ecosystems*. Academic Press, San Diego, 3rd edition, 2001.

D. Wijaya, A. A. Sentosa, and D. W. H. Tjahjo. Phytoplankton community structure and other related eutrophication indications in lake batur, bali-indonesia. *Procedia Environmental Sciences*, 17:736–744, 2013. doi: 10.1016/j.proenv.2013.02.094. URL <https://doi.org/10.1016/j.proenv.2013.02.094>.

Jared D. Willard, Jordan S. Read, Simon Topp, Gretchen J. A. Hansen, and Vipin Kumar. Daily surface temperatures for 185,549 lakes in the conterminous united states estimated using deep learning (1980–2020). *Limnology and Oceanography Letters*, 7(4):287–301, 2022. ISSN 2378-2242. doi: 10.1002/lol2.10249.

P. A. Wiradana, I. K. W. Yudha, and A. T. Mukti. Mass tilapia (*Oreochromis mossambicus*) mortality in floating net cages at batur lake, bangli regency, bali province: a case report. *IOP Conference Series: Earth and Environmental Science*, 1036(1):012068, 2022. doi: 10.1088/1755-1315/1036/1/012068. URL <https://doi.org/10.1088/1755-1315/1036/1/012068>.

N. K. a. L. Wirawan. Viral ribuan ikan mati massal akibat belerang di danau batur, ini penjelasannya, July 2025. URL <https://www.detik.com/bali/berita/d-8010605/viral-ribuan-ikan-mati-massal-akibat-belerang-di-danau-batur-ini-penjasannya>.

Ayi Yustiati and Akira Lusia. A systematic analysis of GIS as a decision support tool for the sustainable management of floating net cages in lake batur, bali, indonesia: Analisis sistematis SIG sebagai alat pendukung keputusan untuk pengelolaan keramba jaring apung yang berkelanjutan di danau batur, bali, indonesia. *Jurnal Teknologi Lingkungan*, 26(2):213–223, July 2025. ISSN 2548-6101. doi: 10.55981/jtl.2025.11386.

Hucai Zhang, Jingan Chen, and Douglas G. Haffner. Plateau lake water quality and eutrophication: Status and challenges. *Water*, 15(2), 2023. ISSN 2073-4441. doi: 10.3390/w15020337. URL <https://www.mdpi.com/2073-4441/15/2/337>.

Senlin Zhu, Bahrudin Hrnjica, Mariusz Ptak, Adam Choiński, and Bellie Sivakumar. Forecasting of water level in multiple temperate lakes using machine learning models. *Journal of Hydrology*, 585:124819, June 2020. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2020.124819.

Senlin Zhu, Ryuichiro Shinohara, Shin-Ichiro S. Matsuzaki, Ayato Kohzu, Mirai Watanabe, Megumi Nakagawa, Fabio Di Nunno, Jiang Sun, Quan Zhou, and Francesco Granata. Diel temperature patterns unveiled: High-frequency monitoring and deep learning in lake kasumigaura. *Ecological Indicators*, 169: 112958, December 2024. ISSN 1470-160X. doi: 10.1016/j.ecolind.2024.112958.