

**PENENTUAN POLA DAN PEMBANGKIT GELOMBANG LAUT DARI  
DATA PUMMA U-TEWS DENGAN PENDEKATAN *DATA ANALYTICS***

**(Skripsi)**

**Oleh**

**RACHMA LINGGA MAULIDYA**

**NPM.2015031006**



**JURUSAN TEKNIK ELEKTRO  
FAKULTAS TEKNIK  
UNIVERSITAS LAMPUNG  
2024**

## ABSTRAK

### **PENENTUAN POLA DAN PEMBANGKIT GELOMBANG LAUT DARI DATA PUMMA U-TEWS DENGAN PENDEKATAN DATA ANALYTICS**

Oleh

**RACHMA LINGGA MAULIDYA**

Sistem peringatan dini tsunami (T-EWS) berbasis *Internet of Things* (IoT) telah dikembangkan dan diterapkan sejak tahun 2019 oleh peneliti dari *Krakatau Research Center* (KRC), bekerja sama dengan PT. Drone Nirwana Bentala (Dronila), serta Unila Robotika dan Otomasi (URO). Sistem yang dibuat menghasilkan data kondisi laut terkini agar dapat mendeteksi dengan cepat tanda-tanda perubahan gelombang yang dapat menyebabkan tsunami. Dari latar belakang tersebut maka diadakan penelitian penerapan *data analytics* sebagai solusi untuk membantu mendeteksi dan mengidentifikasi anomali yang terjadi berdasarkan data kondisi laut yang terekam. Teknik ini memungkinkan untuk mengidentifikasi pola dan karakteristik gelombang serta fluktuasi lainnya dalam data yang diamati. Dimulai dengan pengambilan data prototipe PUMMA U-TEWS dan data cuaca, melakukan *data preprocessing*, mengeksplorasi data, *modelling*, *anomaly detection*, serta melakukan korelasi. Keluaran dari proses ini berupa data yang siap digunakan untuk pemodelan, terdeteksinya anomali data PUMMA U-TEWS, dan teridentifikasinya korelasi antara data PUMMA U-TEWS dengan data cuaca. Pada dataset PUMMA U-TEWS Pulau Panjang terdeteksi sebanyak 20 anomali dan pada dataset PUMMA U-TEWS Desa Gebang terdeteksi sebanyak 23 anomali. Adapun hasil korelasi spearman antara data PUMMA U-TEWS dan data cuaca yaitu terdapat hubungan yang searah dan tidak searah serta nilai koefisien korelasi sedang hingga sangat lemah menandakan data cuaca yang digunakan tidak memiliki pengaruh signifikan dengan tinggi air laut pada data PUMMA U-TEWS. Tinggi air laut dapat saja dipengaruhi oleh faktor lain seperti faktor geologi atau faktor antropogenik.

Kata Kunci: Gelombang Laut, Analisis Data, Deteksi Anomali, Korelasi, Sistem Peringatan Dini Tsunami, PUMMA

## **ABSTRACT**

### **DETERMINING OCEAN WAVE PATTERNS AND GENERATORS FROM PUMMA U-TEWS DATA USING A DATA ANALYTICS APPROACH**

**By**

**RACHMA LINGGA MAULIDYA**

An Internet of Things (IoT)-based tsunami early warning system (T-EWS) has been developed and implemented since 2019 by researchers from Krakatau Research Center (KRC), in collaboration with PT Drone Nirwana Bentala (Dronila), and Unila Robotics and Automation (URO). The system produces data on current sea conditions in order to quickly detect signs of wave changes that can cause tsunamis. From this background, research was conducted on the application of data analytics as a solution to help detect and identify anomalies that occur based on recorded sea condition data. This technique makes it possible to identify patterns and characteristics of waves and other fluctuations in the observed data. It starts with collecting PUMMA U-TEWS prototype data and weather data, performing data preprocessing, exploring data, modeling, anomaly detection, and correlation. The output of this process is data that is ready to be used for modeling, the detection of PUMMA U-TEWS data anomalies, and the identification of correlations between PUMMA U-TEWS data and weather data. In the Panjang Island PUMMA U-TEWS dataset, 20 anomalies were detected and in the Gebang Village PUMMA U-TEWS dataset, 23 anomalies were detected. The results of the Spearman correlation between PUMMA U-TEWS data and weather data, namely there is a unidirectional and unidirectional relationship and the value of the correlation coefficient is moderate to very weak, indicating that the weather data used does not have a significant influence on sea level in the PUMMA U-TEWS data. Sea level can be influenced by other factors such as geological factors or anthropogenic factors.

**Keywords:** Ocean Wave, Data Analytics, Anomaly Detection, Correlation, Tsunami Early Warning System, PUMMA

**PENENTUAN POLA DAN PEMBANGKIT GELOMBANG LAUT DARI  
DATA PUMMA U-TEWS DENGAN PENDEKATAN *DATA ANALYTICS***

**Oleh**

**RACHMA LINGGA MAULIDYA**

**Skripsi**

**Sebagai Salah Satu Syarat untuk Mencapai Gelar  
SARJANA TEKNIK**

**Pada**

**Jurusan Teknik Elektro  
Fakultas Teknik Universitas Lampung**



**FAKULTAS TEKNIK  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

Judul Skripsi : **PENENTUAN POLA DAN PEMBANGKIT  
GELOMBANG LAUT DARI DATA PUMMA  
U-TEWS DENGAN PENDEKATAN DATA  
ANALYTICS**

Nama Mahasiswa : **Rachma Lingga Maulidya**

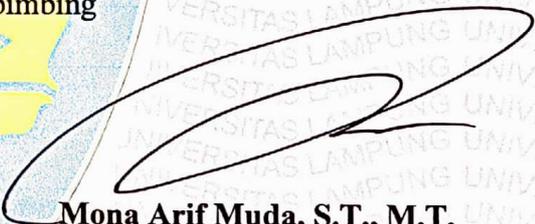
Nomor Pokok Mahasiswa : 2015031006

Program Studi : Teknik Elektro

Fakultas : Teknik



  
**Dr. Ing. Melvi, S.T., M.T.**  
NIP. 19730118 200003 2 001

  
**Mona Arif Muda, S.T., M.T.**  
NIP. 19711112 200003 1 002

2. Mengetahui

Ketua Jurusan  
Teknik Elektro

  
**Herlinawati, S.T., M.T.**  
NIP. 19710314 199903 2 001

Ketua Program Studi  
Teknik Elektro

  
**Sumadi, S.T., M.T.**  
NIP. 19731104 200003 1 001

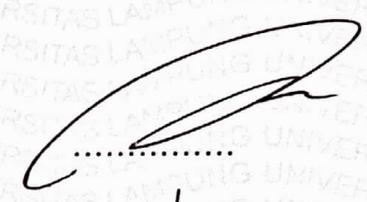
**MENGESAHKAN**

**1. Tim Penguji**

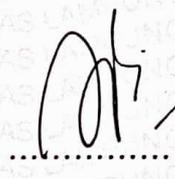
**Ketua : Dr. Ing. Melvi, S.T., M.T.**



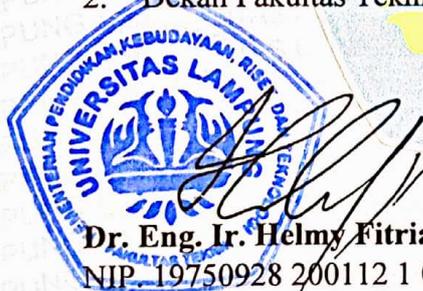
**Sekretaris : Mona Arif Muda, S.T., M.T.**



**Penguji : Dr. Ing. Ardian Ulvan, S.T., M.Sc.**



**2. Dekan Fakultas Teknik**



**Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc.**  
NIP. 19750928 200112 1 002

**Tanggal Lulus Ujian Skripsi : 31 Mei 2024**

## SURAT PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah dilakukan orang lain dan sepanjang pengetahuan saya tidak terdapat atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini sebagaimana yang disebutkan dalam daftar pustaka. Selain itu, saya menyatakan pula bahwa skripsi ini dibuat oleh saya sendiri.

Apabila pernyataan saya tidak benar, maka saya bersedia dikenai sanksi akademik sesuai dengan hukum yang berlaku.

Bandar Lampung, 16 Juli 2024



Rachma Lingga Maulidya  
NPM 2015031006

## RIWAYAT HIDUP



Penulis dilahirkan di Bandar Lampung pada tanggal 30 Mei 2002, sebagai anak kedua dari tiga bersaudara, dari Bapak Supriadi dan Ibu Tuter Lestari.

Pendidikan Sekolah Dasar (SD) diselesaikan di SDN 2 Perumnas Way Kandis pada tahun 2014, Sekolah Menengah Pertama (SMP) di MTs Negeri 2 Bandar Lampung diselesaikan pada tahun 2017, dan Sekolah Menengah Akhir (SMA) di SMA Negeri 9 Bandar Lampung diselesaikan pada tahun 2020.

Tahun 2020, penulis terdaftar sebagai mahasiswa Jurusan Teknik Elektro Universitas Lampung melalui jalur SNMPTN. Selama menjadi mahasiswa penulis pernah menjadi bagian dari Asisten Laboratorium Pengukuran Besaran Listrik dengan memberikan pembimbingan dalam mata kuliah praktikum pengukuran besaran listrik dan praktikum rangkaian listrik. Penulis juga aktif mengikuti Organisasi Himpunan Mahasiswa Teknik Elektro (Himatro) Universitas Lampung sebagai anggota Departemen Pengembangan Keteknikan Divisi Pengabdian Masyarakat tahun 2021 hingga 2022 dan sebagai Wakil Sekretaris Umum tahun 2022 hingga 2023. Penulis mengambil konsentrasi Telekomunikasi dan Teknologi Informasi dan secara aktif mengikuti kegiatan akademik di konsentrasi tersebut. Beberapa kegiatan akademik pada lingkup tersebut yang diikuti oleh penulis diantaranya kegiatan PPK Ormawa di Wilayah Hutan Wisata Mangrove Petengoran, Studi Independen Kampus Merdeka, dan riset dosen AGRI-ICS: *Produk Intelligent Control System Untuk Smart-Greenhouse*.

## PERSEMBAHAN



Dengan segala syukur bagi Allah SWT atas berkah rahmat dan karunia-Nya,  
Kupersembahkan karya ini dengan rasa syukur, hormat dan kasih sayang

### **Kepada Mama dan Papa**

Atas semua rasa kepercayaan, pengorbanan, dukungan doa selama ini  
kepada Rachma sehingga Rachma selalu kuat dan selalu dalam  
lindungan Allah SWT selama menempuh perkuliahan.

### **Kepada Kerabat, Teman, dan Dosen**

Atas semua dukungan, semangat, motivasi, dan ilmu yang diberikan

## **MOTTO HIDUP**

“Jangan pernah remehkan impian, walau setinggi apapun. Sungguh Allah Maha Mendengar. Man Jadda Wajada, Siapa yang bersungguh-sungguh akan berhasil”

(Negeri 5 Menara)

“Berbuat baik, berprasangka baik, ikhlas, dan sabar”

## SANWACANA

Puji Syukur kepada Allah SWT atas berkah, rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian ini dengan judul skripsi “Penentuan Pola dan Pembangkit Gelombang Laut Dari Data PUMMA U-TEWS Dengan Pendekatan *Data Analytics*” dengan baik dan lancar. Skripsi ini disusun untuk memenuhi salah satu syarat dalam memperoleh gelar Sarjana Teknik di Jurusan Teknik Elektro, Fakultas Teknik, Universitas Lampung.

Dalam penyusunan skripsi ini, penulis mendapat banyak dukungan dari berbagai pihak. Untuk itu, penulis mengucapkan terima kasih banyak kepada:

1. Kedua orang tua, Kakak, Adik dan keluarga penulis yang selalu mendo’akan, memberikan semangat, nasihat dan dukungan kepada penulis untuk menyelesaikan skripsi ini dengan lancar.
2. Ibu Dr. Ing. Melvi, S.T., M.T. selaku Dosen Pembimbing Utama yang telah memberikan bimbingan, nasihat, saran, dan motivasi yang membangun kepada penulis sehingga penelitian dan skripsi ini dapat diselesaikan dengan baik.
3. Bapak Mona Arif Muda, S.T., M.T. selaku Dosen Pembimbing Pendamping atas segala waktu yang telah diluangkan untuk memberikan bimbingan, nasihat, arahan dan juga motivasi yang membangun kepada penulis sehingga penelitian dan skripsi ini dapat diselesaikan dengan baik dan lancar.
4. Bapak Dr. Ing. Ardian Ulvan, S.T., M.Sc. selaku Dosen Penguji yang telah meluangkan waktu untuk memberikan arahan, saran dan motivasi kepada penulis sehingga perkuliahan, penelitian dan skripsi ini dapat diselesaikan dengan baik.

5. Bapak Dr. Eng. Ageng Sadnowo Repelianto, S.T., M.T. selaku dosen pembimbing akademik yang memberikan bimbingan dan motivasi yang membangun kepada penulis.
6. Bapak Sumadi, S.T., M.T. selaku Kepala Program Studi Teknik Elektro Universitas Lampung yang telah memberikan dukungan dan kesempatan sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
7. Ibu Herlinawati, S.T., M.T. selaku Ketua Jurusan Teknik Elektro Fakultas Teknik Universitas Lampung.
8. Segenap dosen di Jurusan Teknik Elektro yang telah memberikan ilmu yang bermanfaat kepada penulis selama menempuh pendidikan di perkuliahan.
9. Kakak-kakak Tim Penelitian PUMMA U-TEWS atas segala dukungan, ilmu, masukan, dan waktu yang telah diluangkan, sehingga dapat membantu mempermudah pengerjaan penelitian ini.
10. Tim PPK Ormawa Petengoran yang beranggotakan Tiara Khairunnisa, Muhammad Rafi Ramadhan, Muhammad Rafif Musyaffa, Irfan Mirda, Setiawan Margo Waluyo, Yusuf Riski Sulardi Akbar, dan Muhamad Faris Abdinantara yang telah menjadi orang-orang yang mempercayai penulis selama bekerja dalam satu tim, sebagai teman bertukar pikiran, dan menyemangati penulis untuk bisa menyelesaikan segala hal dengan maksimal.
11. Seluruh teman-teman Telekomunikasi dan Teknologi Informasi 2018, 2019, 2020, 2021 dan 2022 yang telah memberikan dukungan, semangat dan motivasi kepada penulis untuk menyelesaikan skripsi ini.
12. Cahya Andika Salsabilla, Aymanul Fadillah, Adhiva Nur Fadheela, Anita Angraieni, Anna Zakkia Latifah, Anisa Rahmadini, Desi Budiati, Asisten Laboratorium Pengukuran Besaran Listrik, dan Hellios (Teman-teman Angkatan 2020). Terima kasih atas segala proses, pengalaman, semangat dan kebersamaanya selama menempuh pendidikan di Teknik Elektro.
13. Keluarga Besar Himatro yang telah memberikan pengalaman tak terlupakan kepada penulis.
14. Seluruh pihak yang telah membantu penulis menyelesaikan skripsi ini.

Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, maka dari itu penulis mengharapkan kritik dan saran yang membangun dari berbagai pihak. Semoga skripsi ini dapat membantu dan memberikan manfaat bagi penulis dan para pembaca.

Bandar Lampung, Juli 2024

Rachma Lingga Maulidya

## DAFTAR ISI

	<b>Halaman</b>
<b>ABSTRAK</b> .....	<b>ii</b>
<b>ABSTRACT</b> .....	<b>iii</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>vi</b>
<b>SURAT PERNYATAAN</b> .....	<b>vii</b>
<b>RIWAYAT HIDUP</b> .....	<b>viii</b>
<b>PERSEMBAHAN</b> .....	<b>ix</b>
<b>MOTTO HIDUP</b> .....	<b>x</b>
<b>SANWACANA</b> .....	<b>xi</b>
<b>DAFTAR ISI</b> .....	<b>xiv</b>
<b>DAFTAR GAMBAR</b> .....	<b>xvi</b>
<b>DAFTAR TABEL</b> .....	<b>xix</b>
<b>I. PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Tujuan Penelitian.....	2
1.3 Rumusan Masalah .....	2
1.4 Batasan Masalah.....	3
1.5 Manfaat Penelitian.....	3
1.6 Sistematika Penulisan.....	3
<b>II. TINJAUAN PUSTAKA</b> .....	<b>5</b>
2.1 Penelitian Sebelumnya .....	5
2.2 Gelombang Laut .....	12
2.3 <i>Data Analytics</i> .....	13
2.4 <i>Knowledge Discovery in Database (KDD)</i> .....	14
2.5 Pola Data Runtun Waktu ( <i>Time Series</i> ) .....	16
2.6 Peramalan ( <i>Forecasting</i> ).....	17
2.7 <i>Anomaly Detection</i> .....	17
2.8 <i>Anomaly Detection</i> pada <i>Machine Learning</i> .....	18
2.9 Kriteria Anomali .....	20

2.10 <i>Long Short-Term Memory (LSTM)</i> .....	21
2.11 Korelasi Rank Spearman.....	22
2.12 <i>Root Mean Square Error (RMSE)</i> .....	23
2.13 <i>Mean Absolute Percentage Error (MAPE)</i> .....	23
2.14 <i>Mean Absolute Error (MAE)</i> .....	24
2.15 <i>Coefficient of Determination (R-Squared)</i> .....	24
<b>III. METODE PENELITIAN</b> .....	<b>25</b>
3.1 Batu dan Tempat Penelitian.....	25
3.2 Perangkat Penelitian.....	25
3.3 Jenis Penelitian.....	25
3.4 Sumber Data.....	26
3.5 Tahapan Penelitian.....	26
3.6 Rancangan Sistem.....	28
3.7 <i>Anomaly Detection dan Correlation</i> .....	29
<b>IV. HASIL DAN PEMBAHASAN</b> .....	<b>30</b>
4.1 Mengumpulkan Data.....	30
4.2 Skenario Satu.....	32
4.2.1 <i>Data Preprocessing</i> .....	32
4.2.2 <i>Data Exploring</i> .....	39
4.2.3 <i>Modelling</i> .....	44
4.2.4 <i>Anomaly Detection</i> .....	47
4.3 Skenario Dua.....	49
4.3.1 <i>Data Preprocessing</i> .....	50
4.3.2 <i>Data Exploring</i> .....	59
4.3.3 <i>Modelling</i> .....	63
4.3.4 <i>Anomaly Detection</i> .....	66
4.4 <i>Correlation Analysis</i> .....	69
<b>V. SIMPULAN DAN SARAN</b> .....	<b>72</b>
5.1 Kesimpulan.....	72
5.2 Saran.....	73
<b>DAFTAR PUSTAKA</b> .....	<b>74</b>
<b>LAMPIRAN</b> .....	<b>78</b>

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 Ilustrasi Gelombang .....	12
Gambar 2.2 <i>Knowledge Discovery in Databases (KDD) Process</i> .....	15
Gambar 2.3 Pola Data Runtun Waktu.....	17
Gambar 2.4 Prediksi Deret Waktu dengan <i>Prediction Error</i> $e^{(t)}$ . ....	22
Gambar 3.1 Platform PUMMA U-TEWS.....	26
Gambar 3.2 Blok Diagram Penelitian .....	27
Gambar 3.3 Diagram Alir Perancangan dan Implementasi Sistem.....	28
Gambar 4.1 <i>Dashboard website</i> visualcrossing.com .....	31
Gambar 4.2 Proses Penarikan Data.....	31
Gambar 4.3 Penggabungan file csv.....	32
Gambar 4.4 Memanggil Dataset .....	33
Gambar 4.5 Menghapus kolom pada dataset .....	33
Gambar 4.6 Melihat Informasi Dataset.....	34
Gambar 4.7 Penyesuaian tipe data <i>datetime</i> .....	34
Gambar 4.8 Menghapus Nilai Negatif .....	35
Gambar 4.9 Mengecek duplikat data .....	36
Gambar 4.10 Menghapus duplikat dataset PUMMA U-TEWS.....	36
Gambar 4.11 Mengecek <i>range</i> data PUMMA U-TEWS Desa Gebang.....	36
Gambar 4.12 Penggabungan Dataset dengan Dataframe Baru .....	37
Gambar 4.13 Memfilter <i>Range</i> Data.....	38
Gambar 4.14 Pengecekan nilai NaN .....	39
Gambar 4.15 Grafik Dataset PUMMA U-TEWS Pulau Panjang .....	42
Gambar 4.16 Grafik Variabel <i>Temperature</i> .....	42
Gambar 4.17 Grafik Variabel <i>Humidity</i> .....	43
Gambar 4.18 Grafik Variabel <i>Rainfall</i> .....	43

Gambar 4.19 Grafik Variabel <i>Wind Speed</i> .....	43
Gambar 4.20 Grafik Variabel <i>Wind Direction Degrees</i> .....	43
Gambar 4.21 Grafik Variabel <i>Wind Direction</i> .....	43
Gambar 4.22 Data Hasil Prediksi.....	46
Gambar 4.23 Menghitung Residual .....	47
Gambar 4.24 Menentukan ambang batas .....	47
Gambar 4.25 Melihat Jumlah Titik Data Anomali Data PUMMA U-TEWS.	48
Gambar 4.26 Visualisasi Deteksi Anomali Data PUMMA U-TEWS .....	49
Gambar 4.27 Folder File PUMMA U-TEWS Desa Gebang .....	50
Gambar 4.28 Proses Penggabungan File.....	51
Gambar 4.29 Menyimpan Dataset.....	51
Gambar 4.30 <i>Drop</i> Variabel.....	52
Gambar 4.31 Melihat Informasi Dataset.....	52
Gambar 4.32 <i>Error</i> Saat Menyesuaikan Tipe Data <i>Datetime</i> .....	53
Gambar 4.33 Menghapus Nilai Negatif .....	54
Gambar 4.34 Mengecek Duplikat Data.....	55
Gambar 4.35 Menghapus Duplikat Dataset PUMMA Desa Gebang.....	55
Gambar 4.36 Mengecek <i>Range</i> Data PUMMA U-TEWS Desa Gebang.....	55
Gambar 4.37 Penggabungan Dataset Dengan DataFrame Baru .....	56
Gambar 4.38 Menentukan Rentang Data Yang Akan digunakan .....	57
Gambar 4.39 Pengecekan Nilai NaN .....	58
Gambar 4.40 Grafik Dataset PUMMA U-TEWS .....	61
Gambar 4.41 Grafik Variabel <i>Temperature</i> .....	61
Gambar 4.42 Grafik Variabel <i>Humidity</i> .....	61
Gambar 4.43 Grafik Variabel <i>Rainfall</i> .....	62
Gambar 4.44 Grafik Variabel <i>Wind Speed</i> .....	62
Gambar 4.45 Grafik Variabel <i>Wind Direction Degrees</i> .....	62
Gambar 4.46 Grafik Variabel <i>Wind Direction</i> .....	62
Gambar 4.47 Standarisasi Dataset PUMMA U-TEWS Desa Gebang.....	63
Gambar 4.48 Panjang Pembagian Dataset .....	64
Gambar 4.49 Pembagian Timesteps Data Pelatihan .....	64
Gambar 4.50 Membuat Model LSTM .....	65

Gambar 4.51 Konfigurasi Optimizer dan Compile Model.....	65
Gambar 4.52 <i>Training</i> dan <i>Testing</i> Model.....	66
Gambar 4.53 Menghitung Residual .....	67
Gambar 4.54 Menentukan ambang batas Data PUMMA U-TEWS .....	67
Gambar 4.55 Jumlah Titik Anomali Data PUMMA Desa Gebang .....	67
Gambar 4.56 Visualisasi Deteksi Anomali Data PUMMA U-TEWS .....	68
Gambar 4.57 Visualisasi Korelasi PUMMA U-TEWS Desa Gebang .....	70

## DAFTAR TABEL

	Halaman
Tabel 2.1 Penelitian Terdahulu .....	5
Tabel 2.2 Kategori Tinggi Gelombang .....	13
Tabel 2.3 Interpretasi Nilai Rho.....	23
Tabel 4.1 Jumlah Data PUMMA U-TEWS Setelah <i>Preprocessing</i> .....	40
Tabel 4.2 Jumlah Data Cuaca Setelah <i>Preprocessing</i> .....	40
Tabel 4.3 Perhitungan Statistik Deskriptif.....	41
Tabel 4.4 Pembagian data <i>training</i> dan <i>testing</i> PUMMA Pulau Panjang.....	44
Tabel 4.5 Hasil Evaluasi Prediksi Dataset PUMMA Pulau Panjang .....	45
Tabel 4.6 Hasil deteksi anomali PUMMA U-TEWS Pulau Panjang.....	48
Tabel 4.7 Pembagian Metode Penanganan <i>Missing Value</i> .....	58
Tabel 4.8 Jumlah Data PUMMAU-TEWS setelah <i>Preprocessing</i> .....	59
Tabel 4.9 Jumlah Data <i>Microclimate Station</i> setelah <i>Preprocessing</i> .....	59
Tabel 4.10 Perhitungan Statistik Deskriptif.....	60
Tabel 4.11 Hasil Evaluasi Prediksi Dataset PUMMA Desa Gebang.....	63
Tabel 4.12 Hasil deteksi anomali data PUMMA Desa Gebang.....	68
Tabel 4.13 Korelasi PUMMA Gebang terhadap Data <i>Microclimate Station</i> ...	69
Tabel 4.14 Korelasi PUMMA Pulau Panjang terhadap Data Cuaca.....	70

## I. PENDAHULUAN

### 1.1 Latar Belakang

Indonesia memiliki keunggulan di bidang maritim yang tidak dimiliki oleh negara lain. Indonesia merupakan salah satu negara kepulauan terbesar di Asia bahkan di dunia. Hal ini dibuktikan dengan jumlah pulau yang tersebar di 34 provinsi pada tahun 2022 yaitu mencapai 17.001 pulau, dimana enam termasuk pulau besar dan 16.995 pulau lainnya [1]. Secara geografis, Indonesia terletak di antara dua samudera (Samudra Hindia dan Samudera Pasifik), dua benua (Benua Asia dan Benua Australia) dan pada pertemuan dua rangkaian pegunungannya, yaitu Sirkum Pasifik dan Sirkum Mediterania [2]. Berdasarkan kondisi geologisnya, Indonesia merupakan wilayah dengan intensitas bencana yang cukup tinggi seperti gempa bumi dan tsunami. Seluruh bencana yang terjadi disebabkan posisi Indonesia yang dikelilingi oleh tiga lempeng tektonik dunia, yaitu Lempeng Indo-Australia, Lempeng Eurasia, dan Lempeng Pasifik. Posisi Indonesia tersebut terkenal dengan "*Pacific ring of fire*" dimana merupakan negara yang sering dilanda bencana termasuk diantaranya adalah bencana tsunami [3].

Sebagai respons atas potensi bencana yang ada, sistem peringatan dini tsunami (TEWS) berbasis *Internet of Things* (IoT) telah dikembangkan dan diterapkan sejak tahun 2019 oleh peneliti dari *Krakatau Research Center* atau dikenal dengan KRC, bekerja sama dengan PT. Drone Nirwana Bentala (Dronila), serta Unila Robotika dan Otomasi (URO) dimana telah membuat tiga unit PUMMA U-TEWS pada Petengoran Pesawaran, Pulau Panjang (Krakatau), dan juga Desa Canti Lampung. Sistem ini bertujuan untuk membaca kondisi laut terkini serta mendeteksi dengan cepat tanda-tanda awal aktivitas seismik atau perubahan

gelombang yang dapat menyebabkan tsunami, sehingga dapat memberikan peringatan dini kepada masyarakat di sekitarnya. Hasil pembacaan tersebut akan tertampil pada sebuah web dimana nantinya dapat menampilkan data tinggi gelombang sehingga dapat memberikan informasi bahwa akan terjadi tsunami. IDSL (*Inexpensive Device for Sea Level Measurement*) yang dilengkapi dengan sistem peringatan dini tsunami tersebut sebelumnya telah dikembangkan oleh *Joint Reserach Centre – The European Commission* (JRC-EC) dan telah diterapkan di Laut Mediterania [4].

*Data analytics* adalah konsep luas yang berkaitan dengan apa yang terjadi di masa lalu, apa yang terjadi saat ini, dan apa yang diantisipasi dalam menemukan pola, korelasi, dan informasi lainnya. Dalam era modern, *data analytics* telah menjadi alat yang kuat untuk memahami fenomena alam yang kompleks, termasuk gelombang laut. Teknik ini memungkinkan para peneliti untuk mengidentifikasi pola, frekuensi, dan karakteristik gelombang serta fluktuasi lainnya dalam data yang diamati. Dengan menganalisis data PUMMA U-TEWS melalui pendekatan ini, diharapkan dapat diperoleh wawasan mendalam tentang tinggi air laut, melakukan pendeteksian anomali ketinggian muka air laut, serta mengidentifikasi hubungan atau korelasi terhadap variabel lainnya.

## **1.2 Tujuan Penelitian**

Adapun tujuan dari penelitian ini adalah sebagai berikut

1. Mengolah data tinggi air laut dan menampilkan pola tinggi air laut dari data prototipe PUMMA U-TEWS.
2. Melakukan deteksi anomali data ketinggian air laut dan mengidentifikasi hubungan berdasarkan faktor-faktor yang mungkin mempengaruhi pembangkit gelombang laut.

## **1.3 Rumusan Masalah**

Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana menerapkan pendekatan *data analytics* untuk mendeteksi anomali data PUMMA U-TEWS.
2. Apakah hubungan yang diidentifikasi dari pendeteksian anomali data PUMMA U-TEWS berdasarkan faktor-faktor yang mungkin mempengaruhi pembangkit gelombang laut di kawasan tersebut.

#### **1.4 Batasan Masalah**

Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini terbatas pada data PUMMA U-TEWS dalam melakukan deteksi anomali.
2. Analisis ini terbatas pada penggunaan data klimatologi dalam melakukan identifikasi hubungan tinggi air laut berdasarkan anomali yang terdeteksi.
3. Prediksi dan deteksi anomali gelombang laut akan mempertimbangkan keterbatasan data dan metode yang digunakan.
4. Faktor seperti gangguan teknis pada peralatan tidak akan menjadi fokus utama dalam penelitian ini.

#### **1.5 Manfaat Penelitian**

Adapun manfaat penelitian ini adalah dapat mengetahui tentang reliabilitas dan pola data PUMMA U-TEWS, memberikan informasi dalam membuat algoritma prediksi untuk mendeteksi anomali atau pola yang tidak biasa, mengetahui hubungan tinggi air laut terhadap data klimatologi, serta memberikan rekomendasi untuk peningkatan kualitas data dan metode analisis pada penelitian selanjutnya.

#### **1.6 Sistematika Penulisan**

Adapun sistematika penulisan yang digunakan pada penelitian ini untuk memberikan suatu gambaran sederhana mengenai pembahasan tugas akhir serta untuk memudahkan pemahaman materi pada penelitian ini yang dituliskan menjadi beberapa bab, adalah sebagai berikut:

## **I. PENDAHULUAN**

Bab ini menjelaskan tentang latar belakang, tujuan penelitian, rumusan masalah, batasan masalah, manfaat penelitian, hipotesis dan sistematika penulisan pada penelitian ini.

## **II. TINJAUAN PUSTAKA**

Bab ini menjelaskan mengenai teori pendukung yang menjadi pengantar pemahaman dan berkaitan dengan materi penelitian yang diambil dari berbagai sumber ilmiah seperti buku dan jurnal.

## **III. METODOLOGI PENELITIAN**

Bab ini memaparkan metodologi penelitian antara lain waktu dan tempat pengerjaan, alat dan bahan yang digunakan dalam pengerjaan penelitian, serta metode dan diagram penelitian yang akan digunakan dalam pengerjaan penelitian tugas akhir.

## **IV. HASIL DAN PEMBAHASAN**

Bab ini berisikan tentang hasil dan analisa data yang didapatkan dari algoritma dan sistem yang dikembangkan pada penelitian ini.

## **V. PENUTUP**

Bab ini berisikan kesimpulan yang didapat dari hasil analisa dan pembahasan juga berisikan saran yang membangun bagi semua pihak yang terlibat dalam pelaksanaan tugas akhir.

## **DAFTAR PUSTAKA**

Bab ini berisikan referensi dari penulisan dan pelaksanaan proyek.

## II. TINJAUAN PUSTAKA

### 2.1 Penelitian Sebelumnya

Penelitian skripsi ini dibuat dengan memperhatikan analisa dan hasil dari penelitian sebelumnya dalam beberapa tahun terakhir. Terdapat beberapa penelitian terdahulu yang terkait dengan pendekatan *data analytics* dapat dilihat pada tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

No	Nama Peneliti	Judul Penelitian	Hasil Penelitian
1	Ebin Antony. N S Sreekanth. R K Sunil Kumar. Nishanth T. (2021)	<i>Data Preprocessing Techniques for Handling Time Series data for Environmental Science Studies</i>	Artikel ini membahas tentang pentingnya teknik <i>data preprocessing</i> dalam penanganan <i>data time series</i> dalam kajian ilmu lingkungan. Penelitian ini berfokus pada menghilangkan kesalahan dan gangguan pada data sebelum menggunakannya untuk sistem peramalan atau klasifikasi. Artikel ini mengusulkan algoritma <i>Inter Quartile Range (IQR)</i> untuk menangani <i>outlier</i> dan mengeksplorasi berbagai teknik untuk menghilangkan <i>noise</i> , penanganan nilai yang hilang, dan penghalusan. Efektivitas teknik ini

			divalidasi menggunakan data konsentrasi ozon yang dikumpulkan selama periode festival di Kannur, India [5].
2	Akash Singh (2017)	<i>Anomaly detection for temporal data using Long Short-Term Memory (LSTM).</i>	Jurnal ini membahas tentang penggunaan metode deteksi anomali menggunakan <i>LSTM (Long Short-Term Memory)</i> pada data <i>time series</i> dari turbin gas. Penelitian ini menggunakan pendekatan <i>autoencoder</i> untuk <i>features selection</i> dan reduksi, serta model LSTM untuk prediksi. Hasilnya adalah metode tersebut berhasil dalam deteksi anomali pada data <i>time series</i> dari turbin gas dengan kinerja yang baik. Model berhasil mengukur kesalahan agregat untuk mendeteksi perilaku anomali. Metode deteksi menggunakan kesalahan agregat dan kuantil sebagai ambang batas tetap untuk deteksi. Studi tersebut juga mencakup penyetelan <i>hyperparameter</i> dan hasilnya menunjukkan kombinasi terbaik untuk setiap model dan deteksi peristiwa anomali pada sensor-sensor tertentu. Namun, keterbatasan studi ini termasuk kurangnya data pelanggan untuk

			validasi [6].
3	Vamshikrishna Domala. Wonhee Lee. Tae-wan Kim. (2022)	<i>Wave data prediction with optimized machine learning and deep learning techniques.</i>	Artikel penelitian ini berfokus pada prediksi data gelombang menggunakan <i>machine learning</i> yang dioptimalkan dan teknik <i>deep learning</i> . Studi tersebut membandingkan kinerja metode <i>ensemble machine learning</i> seperti <i>random forest</i> , <i>gradient boosting</i> , dan XGBoost, serta metode <i>deep learning</i> FBProphet. Model dievaluasi menggunakan metrik kesalahan, dan hasilnya menunjukkan efektivitas metode <i>ensemble machine learning</i> dan FBProphet dalam memprediksi parameter gelombang [7].
4	Zhang, J. Zhang, H. Ding, S. Zhang, X. (2021)	<i>Power Consumption Predicting and Anomaly Detection Based on Transformer and K-Means</i>	Jurnal ini membahas tentang prediksi konsumsi energi listrik dan deteksi anomali dengan menggunakan model gabungan <i>deep learning</i> dan metode <i>clustering</i> . Model Transformer digunakan untuk memprediksi penggunaan energi listrik pada jam berikutnya, sementara metode <i>clustering</i> K-means digunakan untuk mengoptimalkan hasil prediksi. Anomali kemudian dideteksi dengan membandingkan nilai yang diprediksi dengan nilai test. Hasil pengujian menunjukkan

			<p>bahwa metode yang diusulkan ini memiliki kinerja yang lebih baik daripada model deret waktu LSTM yang umum digunakan [8].</p>
5	Lattawit Kulanuwat (2021)	<p><i>Anomaly Detection Using a Sliding Window Technique and Data Imputation with Machine Learning for Hydrological Time Series</i></p>	<p>Jurnal ini mencakup pengembangan <i>framework</i> untuk deteksi anomali dan pengisian nilai yang hilang dalam data deret waktu hidrologi. Metode yang diusulkan melibatkan penggunaan teknik <i>sliding window</i> dan pembelajaran mesin untuk deteksi anomali, serta berbagai teknik interpolasi untuk pengisian data yang hilang. Selain itu, studi ini juga membandingkan kinerja berbagai metode, seperti median, <i>median absolute deviation</i> (MAD), interpolasi linear, metode spline, dan <i>bidirectional Long Short-Term Memory</i> (LSTM). Hasilnya menunjukkan bahwa metode berbasis median yang dikombinasikan dengan metode interpolasi linear dan <i>spline</i> memberikan hasil yang menjanjikan untuk perilaku data non-siklikal, sementara <i>bidirectional LSTM</i> berperforma baik pada data deret waktu dengan musim yang kuat dan perubahan drastis. Studi ini juga menyoroti</p>

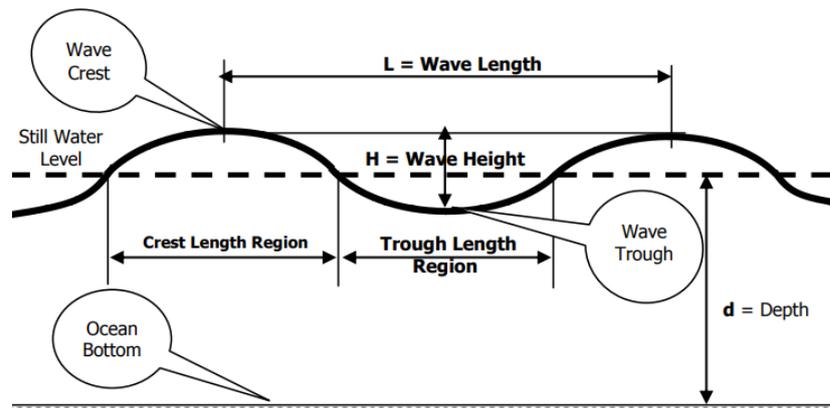
			<p>pentingnya penyetelan parameter, seperti ukuran jendela dan ambang batas, untuk deteksi anomali yang akurat dalam data deret waktu hidrologi. Dengan demikian, hasil dari jurnal tersebut mencakup pengembangan <i>framework</i>, evaluasi kinerja metode deteksi anomali dan pengisian data, serta referensi terkait metode deteksi outlier dalam data deret waktu hidrologi dan lingkungan [9].</p>
6	D A Petrusevich. (2021)	<i>Review of missing values procession methods in time series data</i>	<p>Jurnal ini membahas metode penanganan <i>missing values</i> dalam data <i>time series</i>, termasuk interpolasi, regresi, dan model autoregresif. Metode-metode tersebut diuji pada data <i>time series</i> nilai tukar mata uang, dan hasilnya menunjukkan bahwa interpolasi linear dan <i>cubic splines</i> efektif untuk penanganan nilai yang hilang tunggal maupun berurutan [10].</p>
7	Isack Thomas Nicholaus, Jun Ryeol Park, Kyuol Jung, Jun Seoung Lee, Dae-Ki Kang (2021)	<i>Anomaly Detection of Water Level Using Deep Autoencoder</i>	<p>Studi ini menggunakan teknik <i>deep autoencoder</i> untuk deteksi anomali pada data tingkat air. Penulis mengumpulkan data dari situs sensor di Korea Selatan dan melakukan pra-pemrosesan data termasuk normalisasi, pendekatan <i>sliding window</i>, kompresi data,</p>

			<p>dan pelabelan urutan. Deteksi anomali didasarkan pada kesalahan rekonstruksi dan ambang batas. Evaluasi dilakukan dengan menggunakan akurasi, presisi, recall, F1-score, kurva ROC, dan AUC. Model dilatih dengan API Keras dan optimizer Adam. Hasil menunjukkan kinerja yang menjanjikan dalam mendeteksi anomali dalam data <i>time-series</i>. Studi ini didukung oleh hibah dari pemerintah Korea dan Kementerian Sains, TIK. Data tersedia dengan izin dari penulis dan Infranics Co., Ltd. Tidak ada konflik kepentingan yang diungkapkan. Penelitian ini juga membahas referensi terkait teknik <i>machine learning</i> dan <i>deep learning</i> untuk deteksi anomali dalam berbagai bidang, serta eksperimen yang dilakukan untuk meningkatkan kinerja model dalam mendeteksi anomali pada data sensor. Metode <i>deep autoencoder</i> terbukti efektif dalam mendeteksi anomali dalam data tingkat air [11].</p>
8	González Mata, Alejandro (2020)	<i>A comparison between LSTM and Facebook</i>	Jurnal ini membahas perbandingan antara model LSTM dan model Facebook Prophet dalam konteks

		<i>Prophet models.</i>	peramalan keuangan. Khususnya, penelitian ini fokus pada prediksi harga indeks saham S&P500. Setelah membangun kedua model, terbukti bahwa model LSTM menghasilkan hasil yang lebih baik dan berfungsi sebagai prediktor yang layak dibandingkan dengan strategi perdagangan benchmark lainnya. Model Prophet juga menunjukkan hasil positif dalam investasi, meskipun akurasinya sebagai prediktor tidak sebaik model LSTM [12].
--	--	------------------------	---

## 2.2 Gelombang Laut

Gelombang laut merupakan pergerakan naik dan turun air laut yang disebabkan oleh gaya pembangkitnya. Berdasarkan gaya pembangkitnya, gelombang laut dapat dikategorikan menjadi beberapa jenis antara lain: gelombang laut yang dibangkitkan oleh tiupan angin dipermukaan laut disebut dengan gelombang angin, gelombang laut yang dibangkitkan oleh gaya tarik benda-benda langit terutama matahari dan bulan terhadap bumi disebut gelombang pasang surut, dan gelombang laut yang dibangkitkan oleh letusan gunung berapi atau gempa laut disebut gelombang tsunami, gelombang laut yang dibangkitkan oleh kapal bergerak disebut gelombang kapal dan sebagainya [13]. Gambar 2.1 menunjukkan ilustrasi suatu gelombang.



Gambar 2.1 Ilustrasi Gelombang

(Sumber: <https://www.scribd.com/document/347225123/Teknik-Pantai-pdf>)

Keterangan :

- d = kedalaman laut (jarak antara muka air rerata dan dasar laut)
- H = tinggi gelombang =  $2a$  ( $a = \frac{1}{2} H$ )
- a = amplitudo gelombang
- L = panjang gelombang (jarak antara dua puncak gelombang yang berurutan)
- T = periode gelombang (interval waktu yang diperlukan oleh partikel air untuk kembali pada kedudukan yang sama dengan kedudukan sebelumnya)
- C = kecepatan rambat gelombang =  $L/T$
- $\alpha$  = frekuensi gelombang =  $2\pi/T$
- k = angka gelombang =  $2\pi/L$

Gelombang angin merupakan gelombang yang paling dominan terjadi di permukaan laut dibanding jenis gelombang lainnya, baik ditinjau dari frekuensi kejadian maupun energinya. Angin yang berhembus di atas permukaan laut memindahkan energinya ke air. Kecepatan angin menimbulkan tegangan pada permukaan laut, sehingga permukaan air yang semula tenang akan terganggu dan timbul riak gelombang kecil di atas permukaan laut. Semakin cepat angin, riak tersebut semakin besar dan apabila angin terus berhembus akan terbentuk gelombang. Semakin lama dan semakin kuat angin, semakin besar gelombang laut yang akan terbentuk [14]. Menurut Badan Meteorologi dan Klimatologi, klasifikasi tinggi gelombang laut berdasarkan skalanya ditunjukkan pada tabel 2.1.

Tabel 2.2 Kategori tinggi gelombang

<b>Kategori</b>	<b>Interval (m)</b>
Gelombang Tenang	0.1-0.5 m
Gelombang Rendah	0.5-1.25 m
Gelombang Sedang	1.25-2.50 m
Gelombang Tinggi	2.50-4 m
Gelombang Sangat Tinggi	4-6 m
Gelombang Ekstrem	6-9 m
Gelombang Sangat Ekstrem	9-14 m

Sumber : Badan Meteorologi dan Klimatologi

### **2.3 Data Analytics**

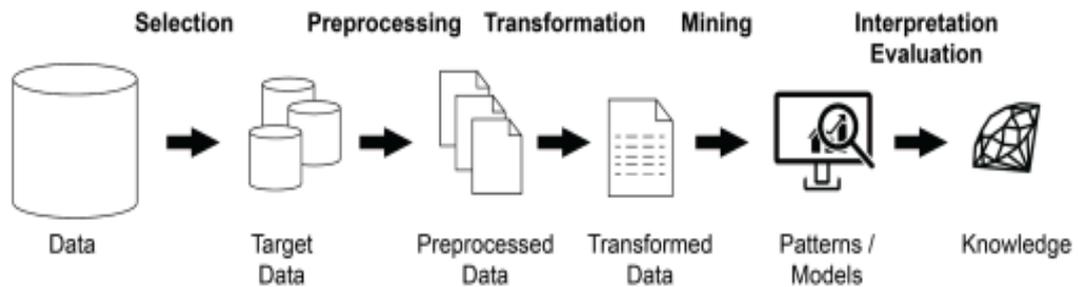
*Data analytics* adalah bidang yang luas, dinamis, dan kompleks karena data hadir dalam berbagai jenis dan berkembang secara signifikan. Selain itu, tujuan analisis bervariasi tergantung pada jenis pengaplikasian yang diperlukan. Oleh karena itu, *data analytics* bertujuan untuk menjawab tiga kategori pertanyaan secara umum yang berkaitan dengan apa yang terjadi di masa lalu, apa yang terjadi saat ini, dan apa yang diantisipasi [15]. Berdasarkan model atau tipe *analytics* nya, *data*

*analytics* terbagi atas beberapa jenis, mulai dari *descriptive analytics*, *predictive analytics*, hingga *prescriptive analytics* [16]. Berikut jenis-jenis *data analytics*:

1. *Descriptive Analytics*, artinya mendeskripsikan semua data sedemikian rupa sehingga dapat dipahami dengan mudah. Sebagian besar skenario mencakup data masa lalu, sekarang, atau historis. Jenis analitik ini digunakan untuk melihat hasil di masa depan berdasarkan data masa lalu. Dalam analitik deskriptif, metode statistik yang digunakan seperti persentase, jumlah, dan rata-rata.
2. *Predictive Analytics* berada di bawah analisis probabilitas yang membantu menentukan prediksi proses di masa depan dan membantu dalam memahami pembuatan pilihan yang akan datang. Saat ini, hal ini menjadi penting bagi organisasi bisnis yang ingin memperoleh keuntungan dalam lingkungan kompetitif ini dengan mendapatkan prediksi mengenai tren masa depan dan menggunakannya untuk membuat keputusan berdasarkan data.
3. *Prescriptive Analytics* adalah jenis analitik lain yang memberikan tindakan sugestif untuk proses pengambilan keputusan. Ini mencakup analisis deskriptif dan prediktif untuk rekomendasi proses pengambilan keputusan.

#### **2.4 Knowledge Discovery in Database (KDD)**

Semua jenis proses *data analytics* termasuk ke dalam proses KDD. *Knowledge Discovery in Database* (KDD) adalah bidang interdisipliner yang berfokus pada metodologi atau teknik untuk mengekstraksi data yang berguna dari sumber. Tantangan dalam mengekstraksi data yang berguna telah menarik minat para peneliti untuk menyelidiki lebih lanjut ekstraksi data di bidang database, pengenalan pola, pembelajaran mesin, visualisasi data, dan komputasi kinerja tinggi. Gambar 2.2 menunjukkan analogi sederhana *Knowledge Discovery in Database* (KDD) *Process* dalam mengekstraksi pengetahuan kepada pengguna. Hal ini menunjukkan bahwa ada lima langkah utama dalam proses KDD; *Selection, Preprocessing, Transformation, Mining, Interpretation/Evaluation* [17]. Berikut adalah tahapan pada *Knowledge Discovery in Databases* (KDD):



Gambar 2.2 *Knowledge Discovery in Databases (KDD) Process*

(Sumber: <https://doi.org/10.1080/17452007.2018.1530092>)

### 1. *Selection*

Pemilihan data berkaitan dengan kebutuhan untuk mengembangkan dan memahami domain aplikasi, menangkap pengetahuan relevan sebelumnya dan mengidentifikasi tujuan proses KDD dari perspektif pengguna akhir.

### 2. *Preprocessing*

*Preprocessing* meliputi pembersihan data dalam hal penanganan data yang hilang, penghapusan duplikat, serta penggabungan dan penyelesaian konflik karena data berasal dari sumber yang heterogen. Persiapan data sangat penting untuk menghasilkan pengetahuan berkualitas tinggi melalui KDD.

### 3. *Transformation*

Transformasi berkaitan dengan reduksi dan proyeksi data dengan tujuan menemukan fitur-fitur yang berguna dan merepresentasikan data sesuai dengan kebutuhan tujuan yang dinyatakan dan algoritma yang dipilih. Hal ini termasuk menemukan representasi data invarian dan menggunakan metode reduksi dimensi untuk mengurangi jumlah efektif variabel yang dipertimbangkan.

### 4. *Mining*

*Data mining* berkaitan dengan mencocokkan tujuan KDD yang ditentukan dengan metode tertentu, misalnya. klasifikasi, regresi, *clustering*, *prediction* atau lainnya. Hal ini mencakup pemilihan algoritme dan metode ekstraksi pola,

serta pertimbangan mengenai kemampuan pengguna akhir untuk menafsirkan model yang dipilih dengan kemampuan prediktif dan akurasi model.

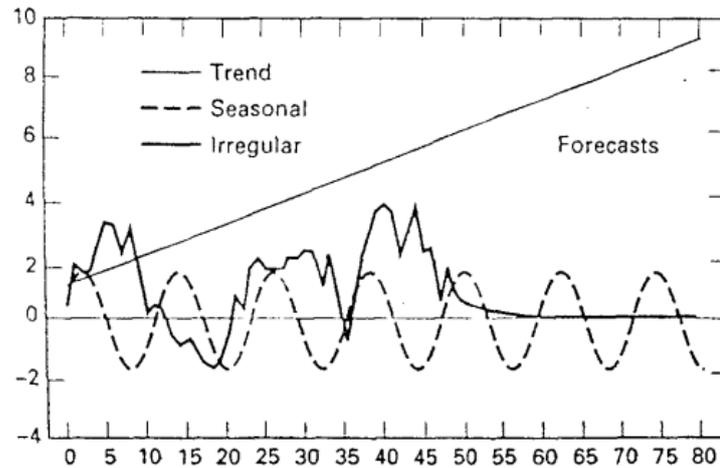
#### 5. *Interpretation/Evaluation*

Langkah terakhir melibatkan interpretasi pola yang dan pemeriksaan validitasnya. Hal ini dapat mencakup visualisasi pola yang ditemukan dan penilaian kegunaannya. Pengetahuan yang ditemukan dalam tahapan ini pula dapat dimanfaatkan, misalnya. mendokumentasikannya, menggunakannya secara langsung, atau mengimplementasikannya ke sistem lain untuk digunakan lebih lanjut [18].

### 2.5 Pola Data Runtun Waktu (*Time Series*)

Dalam memilih suatu metode runtun waktu (*time series*) yang tepat adalah dengan mempertimbangkan jenis pola data [19]. Pola data runtun waktu dapat dibedakan menjadi empat jenis, yaitu seperti pada gambar 2.3.

1. Pola *irregular* terjadi karena data yang diambil tidak dipengaruhi oleh faktor-faktor khusus sehingga pola menjadi tidak menentu dan bersifat acak.
2. Pola *trend* yaitu data cenderung naik maupun turun sepanjang waktu. Tren terjadi ketika ada peningkatan atau penurunan data dalam jangka panjang. Tidak harus linier.
3. Pola *seasonal* (musiman) yaitu data yang dihasilkan oleh kejadian yang terjadi secara musiman atau periodik. Pola *musiman* terjadi ketika rangkaian waktu dipengaruhi oleh faktor musiman seperti waktu dalam setahun atau hari dalam seminggu.
4. Pola siklis terjadi apabila datanya dipengaruhi oleh fluktuasi ekonomi jangka panjang, seperti yang berhubungan dengan siklus perekonomian, biasanya pola ini dipengaruhi oleh faktor eksternal [20].



Gambar 2.3 Pola Data Runtun Waktu

(Sumber: [https://new.mmf.lnu.edu.ua/wp-content/uploads/2018/03/enders\\_applied\\_econometric\\_time\\_series.pdf](https://new.mmf.lnu.edu.ua/wp-content/uploads/2018/03/enders_applied_econometric_time_series.pdf))

## 2.6 Peramalan (*Forecasting*)

Peramalan (*forecasting*) adalah tentang memprediksi tren atau permintaan di masa depan berdasarkan informasi yang tersedia yang mencakup data masa lalu dan pengetahuan tentang kejadian di masa depan yang mungkin berdampak pada perkiraan. Tergantung pada domain aplikasinya, peramalan jangka pendek, jangka menengah, dan jangka panjang dapat dilakukan. Metode peramalan dapat diklasifikasikan menjadi dua yaitu peramalan kualitatif dan peramalan kuantitatif. Jika tidak ada data relevan yang tersedia untuk peramalan, maka metode kualitatif harus digunakan. Di sisi lain, metode kuantitatif digunakan jika data numerik historis tersedia dan ada kemungkinan besar kelanjutan tren masa lalu di masa depan [21].

## 2.7 *Anomaly Detection*

Anomali didefinisikan sebagai pola dalam data yang tidak sesuai atau menyimpang dengan perilaku yang diharapkan atau normal. Anomali merupakan bagian yang sangat kecil dari kumpulan data. Ada tiga jenis anomali yang dipelajari dalam literatur, yaitu anomali titik, anomali sekuensial/kolektif, dan

anomali kontekstual. Masalah dalam menemukan pola seperti itu disebut sebagai *anomaly detection*. Untuk mencapai hasil yang memuaskan dalam deteksi anomali, metode deteksi anomali yang tepat harus dipilih bergantung pada properti data yang diperiksa. Properti berikut ini penting untuk memilih pendekatan yang tepat dalam mendeteksi anomali [22].

1. **Data Temporal vs Non-Temporal:** Data non-temporal dapat berupa gambar medis, rangkaian protein, dll. Data temporal mencakup deret waktu, tetapi juga data dengan stempel waktu dengan interval yang tidak sama.
2. **Data *Univariate* vs *Multivariate*:** Data univariat hanya membutuhkan satu dimensi, misalnya. harga saham, sedangkan data multivariat berisi banyak dimensi. Contoh data multivariat adalah gambar atau rangkaian waktu yang diamati oleh beberapa sensor.
3. ***Labeled* atau *unlabeled data*:** Kumpulan data diberi label jika ada anotasi untuk setiap elemen dalam kumpulan data, yang menentukan apakah titik data tersebut normal atau anomali. Data yang tidak berlabel adalah objek dari metode deteksi anomali tanpa pengawasan.
4. **Jenis anomali dalam kumpulan data:** Anomali titik dideteksi dengan metode klasifikasi langka. Untuk mendeteksi anomali kolektif, metode ini berfokus pada bentuk data yang tidak biasa, sambil mencari penyimpangan. Ini membantu dalam menemukan anomali kontekstual.

## **2.8 *Anomaly Detection* pada *Machine Learning***

*Machine learning* atau pembelajaran mesin merupakan suatu bidang studi cabang dari *Artificial Intelligence* (AI) yang berfokus kepada pengembangan algoritma dan model statistik yang memberi komputer kemampuan untuk belajar dan melakukan tugas-tugas tertentu tanpa diprogram secara eksplisit [23]. Menggunakan *machine learning* dapat membantu dalam mengubah data menjadi informasi atau pola bermakna yang dapat membantu dalam pengambilan keputusan kritis. Teknik dari *machine learning* juga dapat digunakan untuk melakukan pemodelan. Salah satu kasus penggunaan pembelajaran mesin yaitu

deteksi anomali. Ada berbagai jenis metode deteksi anomali dengan pembelajaran mesin, diantaranya *supervised anomaly detection*, *unsupervised anomaly detection*, dan *semi supervised anomaly detection* seperti berikut.

### 1. *Supervised Anomaly Detection*

*Supervised* adalah salah satu pendekatan dalam *machine learning* dimana mesin belajar dari data yang telah diberi label atau label target sebelumnya. Dalam *supervised anomaly detection* ini harus ada set pelatihan untuk objek data dan objek anomali yang diharapkan. Dalam model ini, pendekatannya adalah dengan membangun model prediktif untuk kelas anomali dan normal kemudian membandingkan kedua model tersebut. Namun, dalam mode ini, ada dua masalah yang terjadi. Pertama, jumlah anomali dalam set pelatihan jauh lebih rendah jika dibandingkan dengan kejadian normal. Kedua, label yang tepat dan representatif sulit untuk diidentifikasi, khususnya untuk kelas anomali.

### 2. *Unsupervised Anomaly Detection*

*Unsupervised* adalah pendekatan dalam *machine learning* di mana mesin belajar dari data yang tidak memiliki label atau label target yang diketahui sebelumnya. Kita dapat memberikan skor untuk setiap objek yang menunjukkan sejauh mana anomali tersebut. Kita juga dapat mengamati bahwa jika terdapat banyak anomali yang mirip satu sama lain, maka kita dapat mengelompokkannya sebagai kelompok normal atau skor outliernya rendah. Jadi, kita dapat mengatakan bahwa agar deteksi anomali tanpa pengawasan berhasil, anomali tersebut harus bisa dibedakan [24].

### 3. *Semi Supervised Anomaly Detection*

Terkadang ketika ada data pelatihan dengan objek berlabel normal dan skor diberikan, tetapi tidak memiliki objek anomali, maka kita dapat menerapkan *semi supervised anomaly detection* untuk menemukan anomali tersebut. Deteksi menggunakan objek normal untuk menemukan anomali. Namun, kesulitannya adalah terkadang tidak mudah untuk menemukan kumpulan objek normal yang representatif yang dapat digunakan untuk menemukan anomali [25].

## 2.9 Kriteria Anomali

Dalam mendeteksi anomali *univariate time series* tanpa label data, model mempelajari representasi data yang diberikan dengan cara tanpa pengawasan dengan meminimalkan fungsi tujuan (kerugian) yang ditentukan. Setelah model dilatih, model melakukan diagnosis yang hasilnya berupa skor anomali. Skema untuk memperoleh skor anomali dikelompokkan menjadi tiga jenis, yaitu *reconstruction error*, *prediction error*, dan *dissimilarity* [26].

Terdapat dua cara untuk memperoleh skor anomali dari model prediksi. Pertama menerapkan label biner berdasarkan probabilitas titik data diklasifikasikan sebagai normal. Kesalahan prediksi menunjukkan apakah label yang diharapkan sesuai dengan kebenaran dasar. Pendekatan lainnya adalah memprediksi nilai yang diharapkan untuk langkah waktu berikutnya. Dalam hal ini, kesalahan prediksi adalah sisa antara nilai yang diharapkan dan nilai observasi [26]. Sehingga dalam penelitian ini menggunakan pendekatan kedua untuk menemukan skor anomali yaitu dengan membandingkan hasil prediksi dengan data observasi dan memeriksa perbedaannya dengan menghitung selisih yang disebut sebagai *prediction error* untuk menangkap sejumlah besar titik data anomali menggunakan persamaan 2.1 berikut [27].

$$e^{(t)} = |\hat{y}^{(t)} - y^{(t)}| \quad (2.1)$$

Keterangan:

$e^{(t)}$  : *Prediction Error*

$\hat{y}^{(t)}$  : Data Prediksi

$y^{(t)}$  : Data Aktual

Jika terdapat perbedaan yang cukup besar (*prediction error* lebih tinggi dari ambang batas yang ditentukan), maka dapat diasumsikan bahwa model mengalami kesulitan dalam memprediksi data, sehingga titik data diragukan sebagai anomali. Dalam penelitian ini, perhitungan ambang batas (*threshold*) mengadopsi aturan tiga sigma seperti pada penelitian oleh Isack Thomas Nicholas dkk, yang dinyatakan dalam persamaan (2.2).

$$threshold = mean(e^{(t)}) + 3 \times std(e^{(t)}) \quad (2.2)$$

Dalam Persamaan (2.2), fungsi mean menghitung mean dari kerugian prediksi (yaitu  $e^{(t)}$ ) dan std adalah fungsi deviasi standar yang mengembalikan deviasi standar dari kerugian prediksi.

Lalu, menggunakan persamaan (2.3) untuk memberi label setiap titik data dari prediksi model. Persamaan ini membandingkan kesalahan prediksi,  $e^{(t)}$  terhadap ambang batas yang dihitung dan mengembalikan label untuk rangkaian titik data tertentu  $X_t$ .

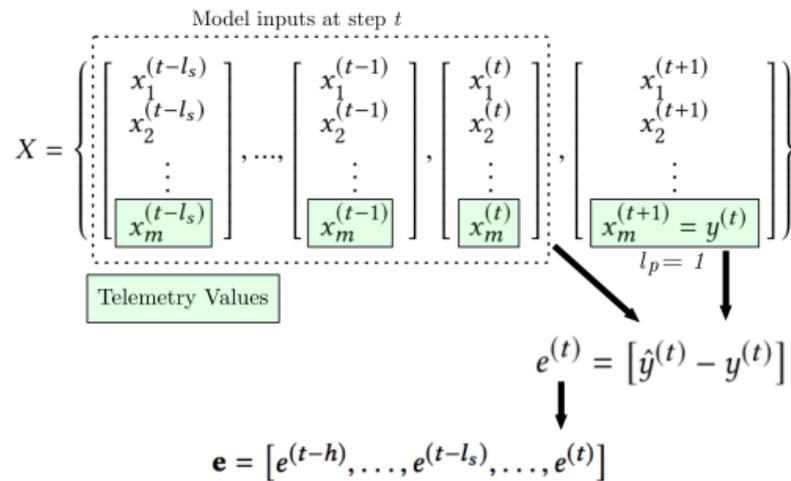
$$l_{x_t} = \theta(e_{x_t}) = \begin{cases} 1 & \text{if } e_{x_t} > threshold \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

Di mana  $l_{x_t}$  adalah label prediksi yang ditetapkan ke urutan  $X_t$  yang bernilai 1 untuk titik data abnormal atau 0 untuk titik data normal. Fungsi  $\theta$  mengembalikan nilai sebagai label setelah menerapkan ambang batas yang dipilih pada kesalahan prediksi [11].

## 2.10 Long Short-Term Memory (LSTM)

*Long Short-Term Memory* (LSTM) adalah perpanjangan dari jaringan saraf berulang (RNN), suatu bentuk jaringan saraf dalam yang terutama digunakan untuk data deret waktu yang diusulkan oleh Hochreiter dan Schmidhuber pada tahun 1997. LSTM terdiri dari beberapa sel memori, masing-masing berisi tiga gerbang: gerbang input, forget, dan output. Gerbang input mengontrol jumlah informasi baru yang masuk ke dalam sel, sedangkan gerbang forget mengontrol jumlah informasi lama yang disimpan di dalam sel. Gerbang keluaran mengontrol jumlah informasi yang dikeluarkan dari sel. Selama setiap timestep, LSTM menerima vektor masukan dan vektor keadaan tersembunyi dari timestep sebelumnya. Arsitektur LSTM dapat ditumpuk untuk membentuk beberapa lapisan, yang memungkinkan jaringan mempelajari pola yang lebih kompleks dalam urutan masukan. Gambar 2.4 merupakan representasi visual dari matriks

masukan yang digunakan untuk prediksi sampai menghasilkan kesalahan prediksi [27].



Gambar 2.4 Prediksi Deret Waktu dengan *Prediction Error*  $e^{(t)}$ .

(Sumber: Hundman K., dkk, 2018)

## 2.11 Korelasi Rank Spearman

Korelasi rank spearman digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel. Sumber data yang digunakan pada analisa korelasi ini tidak harus sama. Korelasi Spearman kuat terhadap outlier dan sangat berguna ketika menangani data ordinal atau tidak terdistribusi normal. Simbol korelasi rank spearman biasanya  $r_s$  atau  $\rho$ . Nilai korelasi rank spearman berada antara  $-1 < \rho < 1$  [28]. Interpretasi nilai  $\rho$  dapat dilihat pada Tabel 2.3 dan rumusnya dapat dilihat pada persamaan 2.4.

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (2.4)$$

Keterangan:

$\rho$  : Koefisien korelasi rank spearman

$d$  : Perbedaan antara peringkat pengamatan berpasangan yang sesuai

$n$  : Banyaknya data (sampel)

Tabel 2.3 Interpretasi Nilai Rho

Rho positif	Rho negatif	Interpretasi
$0,80 \leq \rho \leq 1,00$	$-0,80 \leq \rho \leq -1,00$	Sangat Kuat
$0,60 \leq \rho < 0,80$	$-0,60 \leq \rho < -0,80$	Kuat
$0,40 \leq \rho < 0,60$	$-0,40 \leq \rho < -0,60$	Sedang
$0,20 \leq \rho < 0,40$	$-0,20 \leq \rho < -0,40$	Lemah
$0,00 \leq \rho < 0,20$	$-0 \leq \rho < -0,20$	Sangat Lemah

Sumber: Sugiyono, 2010

### 2.12 Root Mean Square Error (RMSE)

*Root Mean Square Error* (RMSE) merupakan akar kuadrat dari kesalahan rata-rata kuadrat antara nilai prediksi dengan nilai sebenarnya, dimana semakin kecil (mendekati nol), maka hasil prediksi akan semakin akurat. Nilai RMSE dapat dihitung dengan persamaan sebagai berikut [29].

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.5)$$

Keterangan :

- y : nilai aktual
- $\hat{y}$  : nilai hasil prediksi
- n : jumlah data

### 2.13 Mean Absolute Percentage Error (MAPE)

MAPE atau *Mean Absolute Percentage Error* merupakan presentase kesalahan rata-rata antara nilai prediksi dengan nilai sebenarnya. Ini digunakan untuk mengukur keakuratan perkiraan menggunakan rumus di bawah ini.

$$MAPE = \left( \frac{1}{n} \sum \frac{|y_t - \hat{y}_t|}{y_t} \right) \times 100\% \quad (2.6)$$

Keterangan:

- $y$  : nilai aktual  
 $\hat{y}$  : nilai hasil prediksi  
 $n$  : jumlah data

#### 2.14 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) adalah ukuran akurasi ramalan yang paling sederhana. Artinya, rata-rata kesalahan absolut. Kesalahan absolut adalah nilai absolut selisih antara nilai perkiraan dan nilai sebenarnya. MAE memberi tahu kita seberapa besar rata-rata kesalahan yang diharapkan dari perkiraan [29].

$$MAE = \frac{1}{n} \sum |y_t - \hat{y}_t| \quad (2.7)$$

Keterangan :

- $y$  : nilai aktual  
 $\hat{y}$  : nilai hasil prediksi  
 $n$  : jumlah data

#### 2.15 Coefficient of Determination (R-Squared)

R-squared ( $R^2$ ) mewakili koefisien seberapa cocok nilai tersebut dibandingkan dengan nilai aslinya. Nilai dari 0 sampai 1 diartikan sebagai persentase. Semakin tinggi nilainya, semakin baik model tersebut. R-squared dapat dinyatakan dengan:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \quad (2.8)$$

Keterangan :

- $y$  : nilai aktual  
 $\hat{y}$  : nilai hasil prediksi  
 $n$  : jumlah data

### **III. METODE PENELITIAN**

#### **3.1 Waku dan Tempat Penelitian**

Penelitian ini dilakukan sejak disetujuinya proposal penelitian ini yaitu pada bulan Desember 2023 hingga proses pelaporan hasil penelitian ini pada bulan Mei 2024. Penelitian ini dilakukan di Laboratorium Terpadu Teknik Elektro, Jurusan Teknik Elektro, Fakultas Teknik Universitas Lampung.

#### **3.2 Perangkat Penelitian**

Adapun alat dan bahan yang digunakan pada penelitian tugas akhir ini adalah sebagai berikut:

1. Satu unit laptop dengan sistem operasi windows 10 64-bit *intel core i3* sebagai media perancangan algoritma dan seluruh sistem yang dibangun.
2. *Google Sheets* sebagai pengelolaan formula data awal.
3. *Google Collab* dengan spesifikasi GPU NVIDIA-SMI 460.32.03 dan CUDA Version 11.2 dan ram 12.68 GB sebagai tempat untuk *training dataset* dan visualisasi data.
4. Jupyter Notebook version 6.5.4 sebagai tempat untuk analisis data.

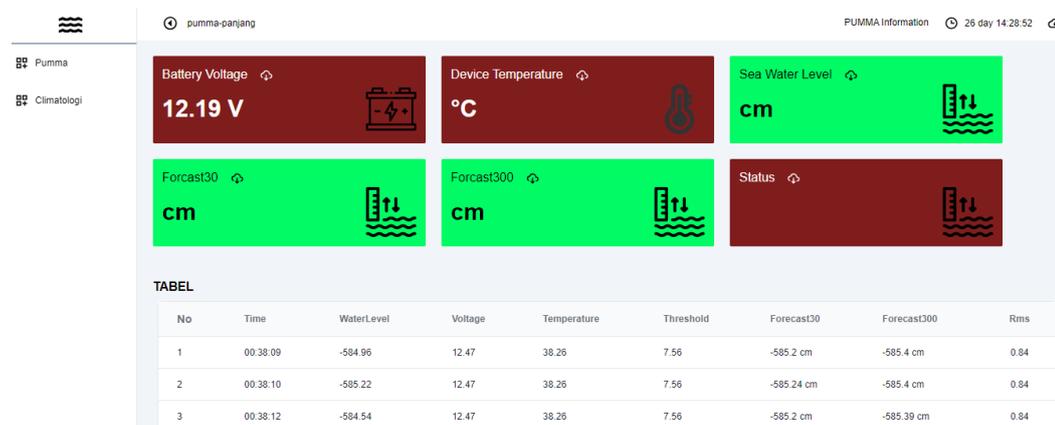
#### **3.3 Jenis Penelitian**

Dalam penelitian ini, kami melakukan identifikasi anomali dengan memanfaatkan metode *data analytics* berbasis *prediction error* lalu melakukan korelasi antar beberapa variabel. Untuk memastikan kerangka kerja dan relevansi penelitian ini, studi literatur dilakukan terlebih dahulu dimana bertujuan untuk memahami riset-

riset sebelumnya yang telah dilakukan terkait dengan *anomaly detection*, serta teknik-teknik komputasional yang telah ada.

### 3.4 Sumber Data

Pada penelitian ini, data yang akan digunakan merupakan data dari prototipe Perangkat Murah Muka Air Laut (PUMMA) milik tim riset dari Universitas Lampung yang berlokasi di Pulau Panjang dan Desa Gebang. Adapun tampilan platformnya dapat dilihat seperti pada gambar 3.1 yang menampilkan data PUMMA U-TEWS Pulau Panjang. Data tersebut memiliki variabel yang akan digunakan dalam penelitian yaitu ketinggian permukaan air laut atau *water level*. Tidak hanya menggunakan data PUMMA U-TEWS, pada penelitian ini juga menggunakan data cuaca Pulau Panjang dan *microclimate station* Desa Gebang. Namun, data yang didapat masih dalam bentuk raw data sehingga perlu dilakukan tahapan *preprocessing data*.

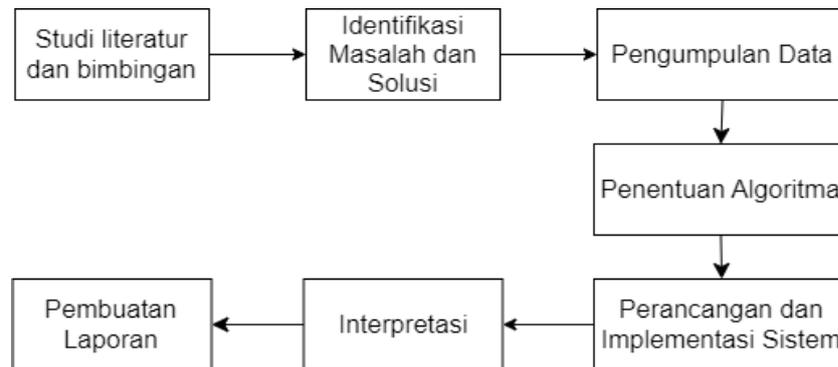


Gambar 3.1 Platform PUMMA U-TEWS

### 3.5 Tahapan Penelitian

Tahapan dan metode yang digunakan dalam penelitian ini menggunakan tahapan yang terstruktur dengan baik seperti halnya dalam proses *data analytics* seperti yang sudah dijelaskan pada gambar 2.2. Langkah-langkah ini dirancang tidak hanya untuk meningkatkan efektivitas dalam pengumpulan dan analisis data,

tetapi juga untuk memastikan efisiensi dalam pemanfaatan sumber daya dan waktu. Pada gambar 3.2 merupakan diagram alur penelitian yang akan dilakukan.



Gambar 3.2 Blok Diagram Penelitian

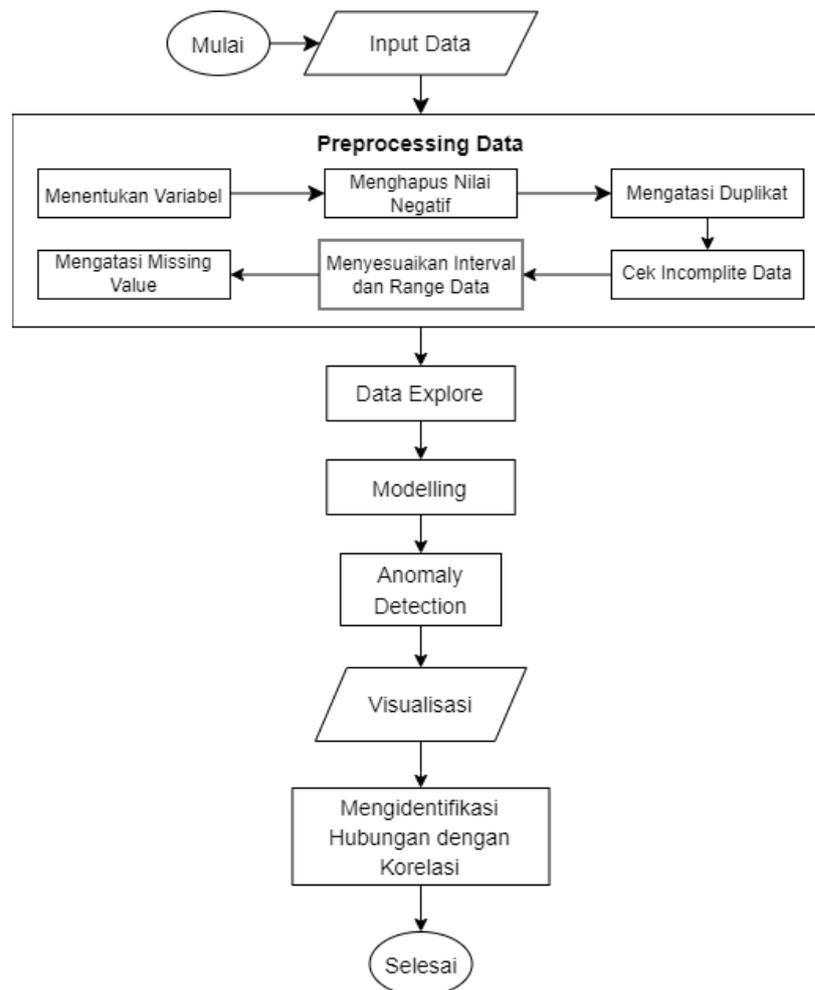
Pada penelitian ini, langkah awal yang dilakukan adalah melakukan studi literatur dan berdiskusi dengan dosen terkait permasalahan yang ada untuk memperoleh wawasan serta temuan sebelumnya. Pada tahap ini dilakukan mencari, mempelajari, serta memahami materi dari berbagai literatur yang berkaitan dengan masalah yang ada dan metode anomaly detection terutama menangani data yang besar. Materi dan literatur yang dipelajari berasal dari berbagai sumber referensi dan sumber-sumber ilmiah lainnya seperti penelitian-penelitian sebelumnya, artikel yang diterbitkan di jurnal dan buku/e-book.

Tahap kedua yaitu mengidentifikasi masalah dari kasus yang dihadapi serta menentukan solusi. Permasalahan kali ini akan berfokus pada identifikasi karakteristik tinggi air laut serta menemukan anomali data dan mencari tahu hubungan anomali tersebut. Berdasarkan pembangkitnya, gelombang laut dapat disebabkan oleh beberapa faktor, seperti angin, arus, kapal, perputaran matahari dan bulan, gempa, dan lainnya.

Selanjutnya, penelitian akan berfokus pada proses pengolahan data. Dimulai dari penentuan algoritma dimana penulis akan mencari algoritma yang relevan untuk data yang digunakan, merancang dan mengimplementasikan sistem deteksi anomali menggunakan algoritma yang telah ditentukan, dan melakukan

interpretasi dalam menyelidiki hubungan tinggi air laut terhadap faktor pembangkit gelombang laut sesuai dengan titik anomali yang teridentifikasi. Setelah semua tahapan pengolahan data selesai, seluruh proses dan eksperimen yang dihasilkan akan dituliskan dan dilaporkan sebagai hasil penelitian dalam pembuatan laporan.

### 3.6 Rancangan Sistem



Gambar 3.3 Diagram Alir Perancangan dan Implementasi Sistem

Pada penelitian ini memiliki beberapa tahapan dalam pelaksanaannya. Hal ini dilakukan untuk menemukan algoritma yang memiliki evaluasi terbaik berdasarkan data yang digunakan serta mendeteksi anomali. Berdasarkan diagram alir pada gambar 3.3, proses pertama yaitu melakukan input data, dimana data

akan di-input, dibaca, dan dilakukan pemeriksaan. Proses kedua yaitu *preprocessing*, dimana dataset yang telah dikumpulkan akan masuk dalam tahap reduksi data, pembersihan atau *cleaning data*, dan penyesuaian. Proses ini sangat dibutuhkan untuk mencari serta mengatasi *missing data*, *correcting invalid values*, *removing duplicate*, dan lain-lain. Proses ketiga adalah *exploring* dengan melihat perhitungan statistik dari data setelah melewati tahapan *preprocessing* dan melakukan visualisasi data. Proses keempat adalah *modelling*. *Modelling* dilakukan dengan membandingkan empat algoritma yaitu prophet, LSTM, ANN, dan neural prophet untuk memperoleh algoritma terbaik yang dapat diterapkan pada data PUMMA U-TEWS dengan melihat hasil evaluasi dari keempat model tersebut, sehingga dapat digunakan untuk proses deteksi anomali nantinya. Evaluasi menggunakan metrik yang sesuai seperti RMSE, MAPE, MAE, dan R-squared. Setelah menemukan model yang relevan, maka proses selanjutnya adalah melakukan deteksi anomali untuk menemukan titik tidak biasa dan memvisualisasikan hasilnya. Kemudian mengidentifikasi apa hubungannya terhadap data cuaca menggunakan metode korelasi seperti korelasi rank spearman.

### **3.7 Anomaly Detection dan Correlation**

Pada tahapan deteksi anomali, algoritma yang relevan dan telah dilatih akan digunakan untuk melakukan prediksi. Data hasil prediksi akan dibandingkan dengan data observasi. Kemudian, menghitung *skor anomali* menggunakan *prediction error* pada persamaan 2.1 dan membandingkannya dengan ambang batas untuk memutuskan apakah suatu *instance* dianggap sebagai anomali serta melakukan korelasi terhadap data klimatologi untuk mengetahui apakah anomali yang terdeteksi memiliki keterkaitan atau hubungan dengan data klimatologi, dimana variabel pada data klimatologi yang akan digunakan seperti *temperature*, *humidity*, *rainfall*, *wind speed*, dan *wind direction*.

## V. SIMPULAN DAN SARAN

### 5.1 Kesimpulan

Kesimpulan yang diperoleh dari penelitian yang telah dilakukan mengenai “PENENTUAN POLA DAN PEMBANGKIT GELOMBANG LAUT DARI DATA PUMMA U-TEWS DENGAN PENDEKATAN *DATA ANALYTICS*” adalah sebagai berikut.

1. Pada penelitian ini telah dibuat dan dirancang algoritma prediksi untuk menentukan banyaknya anomali yang terdeteksi berdasarkan riwayat data dan skenario penelitian, dimana model terbaik untuk melakukan prediksi data PUMMA U-TEWS Pulau Panjang adalah LSTM dengan MAE sebesar 29,74, RMSE sebesar 38,54, MAPE sebesar 31,17%, dan  $R^2$  sebesar 0,80. Kemudian, model terbaik untuk melakukan prediksi data PUMMA U-TEWS Desa Gebang adalah LSTM dengan MAE sebesar 4,02; RMSE sebesar 6,05; MAPE sebesar 5,28%; dan  $R^2$  sebesar 0,93. Visualisasi hasil prediksi data PUMMA U-TEWS Pulau Panjang dan Gebang sudah dapat mengikuti pola data aktual dengan baik.
2. Berdasarkan deteksi anomali yang telah dilakukan pada skenario satu, data PUMMA U-TEWS Pulau Panjang terdeteksi 20 titik anomali, dan pada skenario dua data PUMMA U-TEWS Desa Gebang terdeteksi 23 titik anomali.
3. Berdasarkan hasil analisis korelasi antara anomali PUMMA U-TEWS Desa Gebang dengan data *microclimate station* menghasilkan nilai korelasi kategori sedang hingga lemah dengan korelasi positif (+) antara variabel *water level* dengan *temperature*, *wind direction*, dan *windspeed* menandakan hubungan yang searah, serta korelasi negatif (-) antara variabel *water level* dengan *humidity* dan *rainfall* menandakan hubungan yang tidak searah.

Kemudian, berdasarkan hasil analisis korelasi antara anomali PUMMA U-TEWS Pulau Panjang dengan data cuaca yang bersumber dari *website* menghasilkan nilai korelasi kategori lemah dengan korelasi positif (+) antara variabel *water level* dengan *humidity* dan *wind direction* menandakan hubungan yang searah, serta korelasi negatif (-) antara variabel *water level* dengan *temperature* dan *wind speed* menandakan hubungan yang tidak searah.

4. Berdasarkan hasil korelasi pada penelitian ini didapatkan bahwa anomali tinggi air laut pada data PUMMA U-TEWS tidak memiliki pengaruh signifikan dengan data cuaca dan data *microclimate station*, sehingga cuaca tidak berpengaruh secara signifikan terhadap tinggi air laut, namun bisa saja memiliki pengaruh dengan faktor lain seperti faktor geologi atau faktor antropogenik (aktivitas manusia).

## 5.2 Saran

Saran untuk penelitian selanjutnya berdasarkan penelitian yang telah dilakukan adalah sebagai berikut.

1. Penambahan *dataset* yang relevan untuk meningkatkan tugas identifikasi anomali data tinggi air laut.
2. Mencoba menggunakan metode lain dalam melakukan *anomaly detection* dan analisis sebab akibat.
3. Melakukan analisis korelasi lebih lanjut terkait pengaruh tinggi air laut terhadap faktor-faktor lain yang memungkinkan selain cuaca.

## DAFTAR PUSTAKA

- [1] Data Indonesia. “Indonesia Punya 17.001 Pulau pada 2022, Paling Banyak di Mana?” 2023. [Online]. Available: <https://dataindonesia.id/varia/detail/indonesia-punya-17001-pulau-pada-2022-paling-banyak-di-mana> [Accessed 15/11/2023].
- [2] Susmoro, H. “Penamaan Laut dan Samudera Tata cara dan Implementasinya,” Jakarta: Pusat Hidrografi dan Oseanografi TNI AL. 2019. [Online]. Available: <https://dokumen.tips/documents/penamaan-laut-dan-20200211-penamaan-laut-dan-samudera-tata-cara-dan-implementasinya.html?page=1> [Accessed 15/11/2023].
- [3] W. Warseno et al., “Sustainable Strategy Management Cable-Based Tsunami Detection Equipment Instrumentation,” *In Social, Humanities, and Educational Studies (SHES): Conference Series*, vol. 5, No. 4, pp. 171-179, 2022. [Online]. Available: <https://doi.org/10.20961/shes.v5i4.69044>.
- [4] A. D. G. M. B. Annunziato, “IDSL Sea Level Measurement Devices,” *Technical report JRC-EC*, 2016. doi:10.2788/470647.
- [5] Ebin Antony, N S Sreekanth, R K Sunil Kumar, Nishanth T, "Data Preprocessing Techniques for Handling Time Series data for Environmental Science Studies," *International Journal of Engineering Trends and Technology*, vol. 69, no. 5, pp. 196-207, 2021. Crossref, <https://doi.org/10.14445/22315381/IJETT-V69I5P227>.
- [6] Singh, A. 2017. Anomaly detection for temporal data using long short-term memory (LSTM).
- [7] Vamshikrishna Domala, Wonhee Lee, Tae-wan Kim, “Wave data prediction with optimized machine learning and deep learning techniques,” *Journal of Computational Design and Engineering*, Volume 9, Issue 3, Pages 1107–1122, June 2022. <https://doi.org/10.1093/jcde/qwac048>.
- [8] Zhang J, Zhang H, Ding S and Zhang X, “Power Consumption Predicting and Anomaly Detection Based on Transformer and K-Means,” *Front. Energy Res*, 9:779587, 2021. doi: 10.3389/fenrg.2021.779587.

- [9] Kulanuwat L, Chantrapornchai C, Maleewong M, Wongchaisuwat P, Wimala S, Sarinnapakorn K, Boonya-aroonnet S. "Anomaly Detection Using a Sliding Window Technique and Data Imputation with Machine Learning for Hydrological Time Series," *Water*, 13(13):1862, 2021. <https://doi.org/10.3390/w13131862>.
- [10] Petrusevich, D. A. "Review of missing values procession methods in time series data," *In Journal of Physics: Conference Series*, Vol. 1889, No. 3, p. 032009, April 2021. IOP Publishing. doi: 10.1088/1742-6596/1889/3/032009.
- [11] Nicholaus, Isack Thomas, et al. "Anomaly detection of water level using deep autoencoder," *Sensors*, 21(19):6679, 2021. <https://doi.org/10.3390/s21196679>.
- [12] González Mata, Alejandro. *A comparison between LSTM and Facebook Prophet models: a financial forecasting case study*. 2020.
- [13] Triatmodjo, B. Teknik pantai. 1999. [Online]. Available: <https://www.scribd.com/document/347225123/Teknik-Pantai-pdf>. [Accessed 25/10/2023].
- [14] Wimala L. Dhanista, S.T., M.T., "Gelombang Laut," Institut Teknologi Sepuluh November. 2017. Available: <https://www.its.ac.id/tkelautan/gelombang-laut/>. [Accessed 25/10/23].
- [15] Abdul-Jabbar, S. S. "Data Analytics and Techniques," *ARO-The Scientific Journal of Koya University*, 10(2), 45-55, 2022. <http://doi.org/10.14500/aro.10975>.
- [16] Sharma, R., & Garg, P. "Fundamentals of Data Analytics and Lifecycle". *In Data Analytics and Machine Learning: Navigating the Big Data Landscape* (pp. 19-37), 2024. Singapore: Springer Nature Singapore. [https://doi.org/10.1007/978-981-97-0448-4\\_2](https://doi.org/10.1007/978-981-97-0448-4_2).
- [17] Sabri, I. A. A., Man, M., Bakar, W. A. W. A., & Rose, A. N. M., "Web data extraction approach for deep web using WEIDJ," *Procedia Computer Science*, 163, 417-426, 2019. <https://doi.org/10.1016/j.procs.2019.12.124>.
- [18] Petrova, E., Pauwels, P., Svidt, K., & Jensen, R. L. "Towards data-driven sustainable design: decision support based on knowledge discovery in disparate building data," *Architectural engineering and design management*, 15(5), 334-356, 2019. doi: 10.1080/17452007.2018.1530092.

- [19] Han, J., Kamber, M. and Pei, J. *Data Mining Concepts and Techniques*. 3rd ed. Boston: Elsevier, 2017.
- [20] Enders, W. *Applied econometric time series*. John Wiley & Sons. 2008
- [21] Baban, P. *Validation of Data Streams using Time Series Forecasting*. In GvDB. 2021
- [22] Braei, Mohammad, and Sebastian Wagner. "Anomaly detection in univariate time-series: A survey on the state-of-the-art," *arXiv preprint arXiv:2004.00433*, 2020.
- [23] Retnoningsih, E., & Pramudita, R., "Mengenal Machine Learning Dengan Teknik Supervised Dan Unsupervised Learning Menggunakan Python," *BINA INSANI ICT JOURNAL*, 7(2), 156, 2020. <https://doi.org/10.51211/biict.v7i2.1422>.
- [24] Madhuri, G. Sandhya and Dr. M. Usha. *Anomaly Detection Techniques Causes and Issues*, 2018.
- [25] A. B. Nassif, M. A. Talib, Q. Nasir and F. M. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," in *IEEE Access*, vol. 9, pp. 78658-78700, 2021. doi: 10.1109/ACCESS.2021.3083060.
- [26] Choi, K., Yi, J., Park, C., & Yoon, S. "Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines," *IEEE access*, 9, 120043-120065, 2021. <https://doi.org/10.1109/ACCESS.2021.3107975>.
- [27] Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T., "Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding". In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 387-395), July 2018. <https://doi.org/10.1145/3219819.3219845>.
- [28] Flores, M. J., Ruiz-Vásquez, M., Bastos, A., & Orth, R. "Applying ranking techniques for estimating influence of Earth variables on temperature forecast error". *arXiv preprint arXiv:2403.07966*.
- [29] Febrian, Anggi., et al. "Model Penelusuran Banjir Menggunakan Pendekatan Adaptive Neuro Fuzzy Inference System (ANFIS) (Studi Kasus: Sub DAS Siak)," *Jurnal Online Mahasiswa Fakultas Teknik Universitas Riau*, vol. 2, no. 2, 22 Oct. 2015, pp. 1-9.

- [30] Nizamani, Z., Yih, L. W., Wahab, M. M. A., & Mustaffa, Z, “Determination of correlation for extreme metocean variables,” In *MATEC Web of Conferences* (Vol. 103, p. 04013), 2017, EDP Sciences. <https://doi.org/10.1051/mateconf/201710304013>.
- [31] KANG, Song-Jin and KIM, Chang-Je, “A Study on Shape and Height of Shipwaves,” *Journal of Navigation and Port Research*, vol. 33, no. 2, pp. 105–110, Mar. 2009. <https://doi.org/10.5394/KINPR.2009.33.2.105>.
- [32] Grue, “Ship generated mini-tsunamis,” *Journal of Fluid Mechanics*, vol. 816, pp. 142–166, 2017. doi:10.1017/jfm.2017.67.