

**PREDIKSI JUMLAH SAMPAH BERBASIS LONG SHORT-TERM
MEMORY (LSTM) PADA TEMPAT PEMROSESAN AKHIR
(TPA) BUMIAYU PRINGSEWU**

Skripsi

Oleh

PINKA ANANDA

2115061089



**FAKULTAS TEKNIK
BANDAR LAMPUNG
UNIVERSITAS LAMPUNG
2025**

**PREDIKSI JUMLAH SAMPAH BERBASIS LONG SHORT-TERM
MEMORY (LSTM) PADA TEMPAT PEMROSESAN AKHIR
(TPA) BUMIAYU PRINGSEWU**

**Oleh
PINKA ANANDA**

Skripsi

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

Pada

**Program Studi Teknik Elektro
Jurusan Teknik Elektro
Fakultas Teknik Universitas Lampung**



**FAKULTAS TEKNIK
BANDAR LAMPUNG
UNIVERSITAS LAMPUNG
2025**

ABSTRAK

PREDIKSI JUMLAH SAMPAH BERBASIS LONG SHORT TERM MEMORY (LSTM) PADA TEMPAT PEMROSESAN AKHIR (TPA) BUMIAYU PRINGSEWU

Oleh

PINKA ANANDA

Peningkatan jumlah sampah setiap tahun menjadi tantangan dalam pengelolaan lingkungan, khususnya di Tempat Pemrosesan Akhir (TPA) yang memiliki keterbatasan kapasitas dan daya tampung. Seiring jumlah sampah yang meningkat, diperlukan perencanaan seperti prediksi yang mendukung strategi pengelolaan TPA. Penelitian ini bertujuan untuk memprediksi jumlah sampah di TPA Bumiayu menggunakan model deep learning berbasis Long Short Term Memory yang memiliki kemampuan menangkap pola temporal jangka panjang. Penelitian ini menggunakan tiga arsitektur LSTM yaitu Single LSTM, Double LSTM (Stacked), dan Bidirectional LSTM untuk mendapatkan akurasi terbaik serta mengukur pengaruh fitur eksternal seperti data cuaca, penduduk, dan ekonomi. Data historis 2021–2024 digunakan untuk pelatihan model sedangkan prediksi dilakukan untuk periode 2025–2030. Hasil evaluasi menunjukkan bahwa arsitektur single LSTM menghasilkan performa terbaik dengan MAPE 6,37%, MAE 7,0278, RMSE 9,3872, dan *failure rate* sebesar 6,64%. Hasil prediksi jangka panjang menunjukkan tren peningkatan jumlah sampah secara bertahap yang dapat menjadi strategi dalam perencanaan pengelolaan sampah berkelanjutan.

Kata Kunci: Long Short Term Memory (LSTM), Prediksi Jumlah Sampah, Peramalan deret waktu, Fitur Eksternal

ABSTRACT

PREDICTION OF THE AMOUNT OF WASTE BASED ON LONG SHORT TERM MEMORY (LSTM) AT THE BUMIAYU PRINGSEWU FINAL PROCESSING STATE (TPA)

By

PINKA ANANDA

The increasing amount of waste each year poses a challenge in environmental management, particularly at Final Waste Processing Sites (FWPS) that have limited capacity and storage space. As the amount of waste increases, planning such as predictions is needed to support FWPS management strategies. This study aims to predict the amount of waste at the Bumiayu FWPS using a deep learning model based on Long Short Term Memory (LSTM) that has the ability to capture long-term temporal patterns. This study employs three LSTM architectures—Single LSTM, Double LSTM (Stacked), and Bidirectional LSTM—to achieve the best accuracy and assess the impact of external features such as weather data, population, and economic indicators. Historical data from 2021 to 2024 was used for model training, while predictions were made for the period from 2025 to 2030. Evaluation results show that the Single LSTM architecture achieves the best performance with a MAPE of 6.37%, MAE of 7.0278, RMSE of 9.3872, and a failure rate of 6.64%. Long-term prediction results indicate a gradual increase in waste volume, which can serve as a strategy for sustainable waste management planning.

Keywords: Long Short Term Memory (LSTM), Waste generation prediction, Time series forecasting, External feature

Judul Skripsi

**: PREDIKSI JUMLAH SAMPAH BERBASIS
LONG SHORT TERM MEMORY (LSTM)
PADA TEMPAT PEMROSESAN AKHIR
(TPA) BUMIAYU PRINGSEWU**

Nama Mahasiswa

: Pinka Ananda

Nomor Pokok Mahasiswa

: 2115061089

Program Studi

: Teknik Informatika

Fakultas

: Teknik



1. Komisi Pembimbing

Pembimbing Utama

Pembimbing Pendamping

Ir. M. Komarudin, S.T., M.T.

NIP. 196812071997031006

Yessi Mulyani, S.T., M.T.

NIP. 197312262000122001

2. Mengetahui

Ketua Jurusan

Ketua Program Studi

Herlinawati, S.T., M.T.

NIP. 197103141999032001

Yessi Mulyani, S.T., M.T.

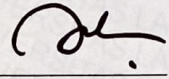
NIP. 197312262000122001

MENGESAHKAN

1. Tim Penguji

Ketua

: **Ir. M. Komarudin, S.T., M.T.**



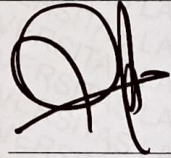
Sekretaris

: **Yessi Mulyani, S.T., M.T.**



Anggota

: **Puput Budi Wintoro, S.Kom., M.T.I.**



2. Dekan Fakultas Teknik



Dr. Eng. Ir. Helmy Fitriawan, S.T., M.Sc.

NIP. 19759282001121002

Tanggal Lulus Ujian Skripsi : 31 Juli 2025

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan bahwa skripsi dengan judul “Prediksi Jumlah Sampah Berbasis Long Short Term Memory (LSTM) Pada Tempat Pemrosesan Akhir (TPA) Bumiayu Pringsewu” adalah hasil karya saya sendiri. Segala hasil yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila di kemudian hari terbukti bahwa pernyataan ini tidak benar, maka saya bersedia menerima sanksi sesuai dengan ketentuan hukum atau akademik yang berlaku.

Bandar Lampung, 20 Agustus 2025

Pembuat Pernyataan,



Pinka Ananda

NPM. 2115061089

RIWAYAT HIDUP



Penulis lahir pada 3 Desember 2002 di Pajaresuk. Penulis adalah anak kedua dari Bapak Paino dan Ibu Holnida. Penulis memulai pendidikan sekolah dasar di SD Negeri 1 Pajaresuk dan lulus pada tahun 2015, kemudian menyelesaikan pendidikan menengah pertama di SMP Negeri 1 Pringsewu dan lulus pada tahun 2018 lalu melanjutkan pendidikan menengah atas di SMA Negeri 1 Pringsewu. Pada tahun 2021, penulis menyelesaikan sekolah menengah atas diterima sebagai mahasiswa baru Teknik Informatika Fakultas Teknik di Universitas Lampung melalui jalur SBMPTN. Selama kuliah penulis mengikuti beberapa kegiatan seperti

1. Menjadi anggota Himpunan Mahasiswa Teknik Elektro Universitas Lampung di Departemen Pengembangan Keteknikan pada divisi Pengabdian Masyarakat periode 2022/2023 dan divisi Penelitian dan Pengembangan periode 2023/2024.
2. Mengikuti program Studi Independen Kampus Merdeka dari Kementerian Pendidikan dan Budaya pada tahun 2023 pada mitra Startup Campus dengan kelas Artificial Intelligence (Computer Vision).
3. Mengikuti program Kuliah Kerja Nyata di desa Pulo Gadung, Kecamatan Penawar Tama, Kabupaten Tulang Bawang, Provinsi Lampung pada bulan Januari sampai dengan Februari tahun 2024.
4. Mengikuti program Magang Bersertifikat Kampus Merdeka Batch VII di Dinas Sosial Kota Medan sebagai Analis Data sejak Februari sampai dengan Juni 2024

MOTTO

“Dan orang-orang yang berjihad (mencari keridhaan) Kami, benar-benar akan Kami tunjukkan kepada mereka jalan-jalan Kami. Sesungguhnya Allah benar-benar beserta orang-orang yang berbuat baik”

(QS. Al-Ankabut: 69)

“The best way to predict the future is to invent it.”

(Alan Kay – Author of The Early History of SmallTalk)

“Success is not final, failure is not fatal: it is the courage to continue the counts”

(Winston Churchill)

PERSEMBAHAN

Bismillahirrahmanirrahim segala puji bagi Allah SWT yang telah melimpahkan nikmat, rahmat, hidayat, dan kekuatan sehingga skripsi ini dapat terselesaikan.

Dengan segenap rasa syukur skripsi ini kupersembahkan untuk kedua orang tua tercinta bapak Paino dan ibu Holnida. Terima kasih untuk setiap doa yang tidak pernah putus, peluh yang tak terlihat, serta semangat yang selalu menguatkan. Terima kasih atas segala pengorbanan dan kerja keras yang tak ternilai. Terima kasih atas dukungan dalam setiap langkah yang saya pilih. Terima kasih juga untuk kakak saya Renal Gusnanda yang turut selalu mendukung saya. Besar harapan saya agar orang tua dan kakak saya selalu sehat, panjang umur, dan melihat keberhasilan lain yang akan penulis raih.

Persembahan ini juga diucapkan untuk diri sendiri yang dapat menyelesaikan perjalanan perkuliahan dengan segala ilmu dan pengalaman yang didapat. Terima kasih sudah berusaha, bertahan, dan tidak menyerah meskipun harus melalui kecemasan yang kadang muncul. Semoga dengan selesainya skripsi ini dapat menjadi langkah baru dalam mewujudkan cita-cita yang lebih besar, menjadi pribadi yang dapat bermanfaat, yang dapat menebar ilmu dan kebaikan.

SANWACANA

Alhamdulillahil'alamin puji syukur penulis panjatkan kepada Allah SWT atas berkat rahmatnya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Prediksi Jumlah Sampah Berbasis Long Short Term (LSTM) Pada Tempat Pemrosesan Akhir (TPA) Bumiayu Pringsewu". Skripsi ini dapat terselesaikan dengan dukungan, doa, dan peran dari berbagai pihak. Dengan penuh rasa hormat dan syukur, penulis mengucapkan terima kasih sebesar-besarnya kepada:

1. Orang tua tercinta yang selalu memberikan doa, dukungan baik emosional maupun material, serta semangat yang tidak pernah henti sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
2. Bapak Dr. Eng Helmy Fitriawan, S.T., M. Sc. selaku Dekan Fakultas Teknik Universitas Lampung
3. Ibu Herlinawati, S.T., M.T. selaku Ketua Jurusan Teknik Elektro Universitas Lampung
4. Ibu Yessi Mulyani, S.T., M.T. selaku Ketua Program Studi Teknik Informatika Universitas Lampung dan selaku dosen pembimbing pendamping yang telah memberikan dukungan, bimbingan, masukan, dan arahan baik saat masa perkuliahan maupun dalam menyelesaikan skripsi ini.
5. Bapak Ir. M. Komarudin, S.T., M.T. selaku dosen pembimbing utama yang selalu memberikan waktunya untuk memberikan arahan dan bimbingan selama proses pengerjaan penelitian sampai dengan skripsi ini selesai. Terima kasih ilmu dan nasihat berharga yang telah diberikan.
6. Bapak Puput Budi Wintoro, S.Kom., M.T.I. selaku dosen penguji yang telah memberikan saran serta masukan dalam menyempurnakan skripsi ini.

7. Bapak Rio Ariestia Pradipta selaku dosen pembimbing akademik yang sudah membantu dalam perjalanan perkuliahan ini. Terima kasih untuk arahan, saran, dan bimbingannya.
8. Para dosen, civitas akademik di Program Studi Teknik Informatika Universitas Lampung.
9. Kakak laki-laki saya Renal Gusnanda yang telah memberikan dukungan dan semangat bagi penulis.
10. Teman-teman saya Chandrika, Syifa, Reginia, Murti, Vania, Hilda, Annisya, dan Tyas yang dari awal semester satu telah menemani perjalanan perkuliahan ini menjadi tim sukses, menjadi tempat bertukar pikiran, berkeluh kesah, dan memberikan semangat serta dukungan.
11. Terima kasih untuk teman-teman PSTI angkatan 21 yang tidak dapat penulis sebutkan satu persatu atas bantuan, semangat, dan cerita-cerita yang sudah dilalui bersama selama masa perkuliahan.
12. Pihak Dinas Lingkungan Hidup Kabupaten Pringsewu yang telah mengizinkan penulis dalam melakukan penelitian.

Akhir kata, penulis berharap semoga hasil penelitian ini dapat memberikan manfaat baik bagi perkembangan ilmu di bidang terkait maupun juga pembaca yang membaca hasil penelitian ini.

Bandar Lampung, 17 Agustus 2025
Penulis,

Pinka Ananda

DAFTAR ISI

Halaman

DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR.....	xvi
I. PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	6
1.6 Sistematika Penulisan.....	6
II. TINJAUAN PUSTAKA.....	8
2.1 Sampah.....	8
2.2 TPAS Bumi Ayu.....	10
2.3 Data Berkala (Time Series).....	10
2.4 Artificial Intelligence.....	10
2.5 <i>Machine Learning</i>	12
2.6 Deep Learning.....	13
2.6.1 Jaringan Saraf Tiruan.....	14
2.6.2 Recurrent Neural Network (RNN).....	18
2.6.3 Long Short-Term Memory (LSTM).....	21
2.7 Fungsi Aktivasi.....	27
2.8 Backpropagation.....	29
2.9 CRISP-DM.....	31
2.10 Evaluasi Model.....	33
2.11 Dekomposisi.....	35
2.12 Korelasi Pearson.....	35
2.13 Interquartile Range (IQR).....	36
2.14 Microsoft Excel.....	37
2.15 Google Colaboratory.....	37

2.16 Streamlit.....	38
2.17 Python.....	38
2.18 Penelitian Terkait.....	39
III. METODE PENELITIAN.....	46
3.1 Waktu dan Tempat Penelitian.....	46
3.2 Alat dan Bahan Penelitian.....	47
3.2.1 Alat.....	47
3.2.2 Bahan.....	48
3.3 Tahapan Penelitian.....	48
3.3.1 Tahap Pemahaman Bisnis.....	50
3.3.2 Tahap Pemahaman Data.....	53
3.3.3 Tahap Persiapan Data.....	56
3.3.4 Tahap Pemodelan.....	57
3.3.5 Tahap Evaluasi.....	57
3.3.6 Tahap Penyebaran (Deployment).....	57
IV. HASIL DAN PEMBAHASAN.....	59
4.1 Hasil Penelitian.....	59
4.1.1 Pemahaman Data.....	59
4.1.2 Persiapan Data.....	86
4.1.3 Pemodelan.....	100
4.1.4 Evaluasi.....	105
4.1.5 Deployment / Prediksi.....	114
4.2 Pembahasan Penelitian.....	118
V. PENUTUP.....	122
5.1 Kesimpulan.....	122
5.2 Saran.....	123
DAFTAR PUSTAKA.....	125

DAFTAR TABEL

	Halaman
Tabel 2.1 Tingkat Akurasi Berdasarkan Nilai MAPE	33
Tabel 2.2 Penelitian Terkait	38
Tabel 3.1 Jadwal dan Kegiatan Penelitian	46
Tabel 3.2 Perangkat Lunak	47
Tabel 3.3 Sampel Data Jumlah Sampah TPA Bumiayu Tahun 2021-2024	53
Tabel 3.4 Data Jumlah Penduduk dan PDRB Per Kapita Kabupaten Pringsewu	54
Tabel 3.5 Sample Data Cuaca Kabupaten Pringsewu	54
Tabel 4.1 Data Penelitian	58
Tabel 4.2 Data Sampah Harian	59
Tabel 4.3 Rekayasa Fitur Waktu	90
Tabel 4.4 Hasil Evaluasi Kombinasi Arsitektur dan Hyperparameter	104
Tabel 4.5 Evaluasi Fitur Eksternal	112

DAFTAR GAMBAR

	Halaman
Gambar 1.1 Jumlah Sampah TPA Bumiayu Per Bulan (2021-2024)	2
Gambar 2.1 Penerapan Konsep Kecerdasan Buatan	11
Gambar 2.2 Kategori Machine Learning	12
Gambar 2.3 Jaringan Umpan Maju Lapis Tunggal	15
Gambar 2.4 Jaringan Umpan Maju Multilapis	16
Gambar 2.5 Jaringan Berulang Lapis Tunggal	17
Gambar 2.6 Jaringan Berulang Multilapis	17
Gambar 2.7 Arsitektur Umum RNN	19
Gambar 2.8 RNN dalam banyak layer	19
Gambar 2.9 Arsitektur RNN	20
Gambar 2.11 Forget Gate	22
Gambar 2.12 Input Gate	23
Gambar 2.13 Cell State	23
Gambar 2.14 Output Cell	24
Gambar 2.15 Single Layer LSTM	24
Gambar 2.16 Double Layer LSTM	25
Gambar 2.17 Bidirectional LSTM	26
Gambar 2.21 Arsitektur Backpropagation	29
Gambar 2.22 Tahapan CRISP-DM	31
Gambar 3.1 Flowchart Tahapan Penelitian	50
Gambar 4.1 Analisis Statistik Data Sampah Harian	61
Gambar 4.2 Missing Values Data Sampah	62
Gambar 4.3 Outlier Data Sampah Harian	63
Gambar 4.4 Grafik Dekomposisi Data Sampah (365 hari)	65
Gambar 4.5 Grafik Dekomposisi Data Sampah (30 hari)	66
Gambar 4.6 Rata-rata Volume Sampah per Hari	67
Gambar 4.7 Rata-Rata Volume Sampah Hari Kerja vs Akhir Pekan	68
Gambar 4.8 Rata-rata Volume Sampah per Bulan	69
Gambar 4.9 Distribusi Total Volume Sampah (m ³)	70

Gambar 4.10 Distribusi Lonjakan Total Volume Sampah (m ³)	70
Gambar 4.11 Data Cuaca Harian	71
Gambar 4.12 Analisis Statistik Data Cuaca Harian	73
Gambar 4.13 Missing Values Data Cuaca Harian	74
Gambar 4.14 Outlier Data Cuaca Harian	75
Gambar 4.15 Boxplot Outlier Data Cuaca	77
Gambar 4.16 Heatmap Korelasi Fitur Cuaca vs Total Volume Sampah	78
Gambar 4.17 Relasi antara Fitur Cuaca dan Total Volume Sampah	80
Gambar 4.18 Data Sosial Ekonomi	82
Gambar 4.19 Resampling Data Sosial Ekonomi	83
Gambar 4.20 Analisis Statistik Data Sosial Ekonomi	83
Gambar 4.21 Missing Values Data Sosial Ekonomi	84
Gambar 4.22 Heatmap Korelasi Data Sosial Ekonomi dan Total Volume Sampah	84
Gambar 4.23 Pembersihan Missing Values Data Sampah	85
Gambar 4.24 Hasil Pembersihan Missing Values Data Sampah	86
Gambar 4.25 Data Sampah Harian	87
Gambar 4.26 Hasil Pembersihan Outlier Data Cuaca	88
Gambar 4.27 Fungsi Penanganan Outlier Data Cuaca	88
Gambar 4.28 Penggabungan Dataset Lengkap	93
Gambar 4.29 Penggabungan Dataset Spesifik	94
Gambar 4.30 Hasil Sliding Window	95
Gambar 4.31 Hasil Pembagian Data	96
Gambar 4.32 Nilai Asli Sebelum Normalisasi (Fitur)	97
Gambar 4.33 Nilai Setelah Normalisasi (Fitur)	98
Gambar 4.34 Target Sebelum dan Sesudah Normalisasi	98
Gambar 4.35 Fungsi Pemanggilan Bangun Model	99
Gambar 4.36 Fungsi Pelatihan Model	100
Gambar 4.37 Grafik Training dan Validation Loss Single Layer Stabil	101
Gambar 4.38 Grafik Training dan Validation Loss Double Layer Stabil	102
Gambar 4.39 Grafik Training dan Validation Loss Bidirectional Layer Stabil	103
Gambar 4.40 Grafik Aktual dan Prediksi Model Terbaik	108
Gambar 4.41 Grafik Aktual dan Prediksi Model Terbaik (Zoom in)	109
Gambar 4.42 Seleksi Fitur Mutual Information Fitur Cuaca	112
Gambar 4.43 Hasil Prediksi Total Sampah 2025–2030	114
Gambar 4.44 Dashboard Prediksi Jumlah Sampah TPA Bumiayu	116
Gambar 4.45 Grafik Aktual Prediksi Model Prophet	118
Gambar 4.46 Evaluasi Tambahan Data Aktual (2025) dan Data Prediksi (2025)	119

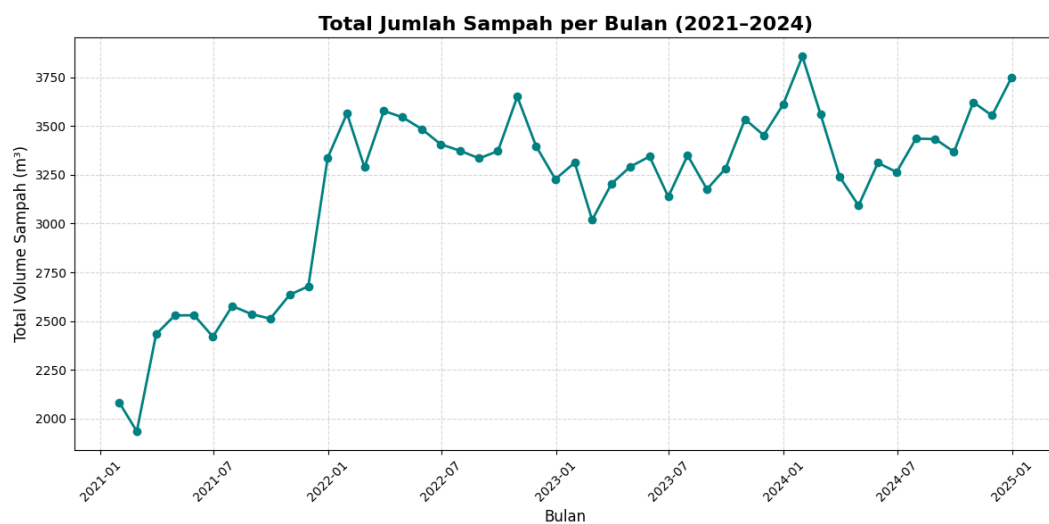
I. PENDAHULUAN

1.1 Latar Belakang

Pengelolaan sampah menjadi salah satu tantangan lingkungan terbesar di berbagai kota di Indonesia. Setiap tahun, Jumlah timbulan sampah yang terus meningkat karena adanya perubahan seperti populasi yang meningkat, cuaca yang berubah-ubah dan aktivitas ekonomi yang lebih aktif. Menurut data jumlah penduduk Indonesia per 30 Juni tahun 2024 yang dikutip dari Direktorat Jenderal Kependudukan dan Pencatatan Sipil (Ditjen Dukcapil) Kementerian Dalam Negeri (Kemendagri) merilis Data Kependudukan Bersih (DKB) tercatat mencapai 280.477.584 jiwa dan menduduki nomor empat sebagai jumlah penduduk terbanyak di dunia dan jumlah ini naik sebanyak 1.752.156 jiwa dibanding tahun sebelumnya [1]. Sedangkan berdasarkan data Sistem Informasi Pengelolaan Sampah Nasional (SIPSN), timbulan sampah nasional pada 2023 yang dihasilkan Indonesia mencapai 39,737,086.46 ton timbulan sampah [2].

Jumlah sampah yang terus meningkat setiap tahun tentu menjadi suatu tantangan bagi pengelola lingkungan di berbagai daerah, terutama dalam keterbatasan kapasitas lahan di tempat pemrosesan akhir (TPA) yang selama ini menjadi solusi dalam penanganan sampah di berbagai daerah. Pengelolaan sampah di berbagai daerah yang masih bersifat reaktif belum adanya perencanaan berbasis prediktif. Akibatnya TPA berisiko mengalami kelebihan kapasitas yang jika tidak diatasi dengan baik, maka dapat berkontribusi dalam memberikan dampak terhadap pencemaran lingkungan sehingga berisiko menghambat komitmen Indonesia untuk mencapai target *Net-Zero Emission* pada tahun 2060 yang salah satunya dapat dicapai dengan pengelolaan limbah termasuk sampah.

TPA Bumiayu adalah salah satu fasilitas pengelolaan sampah di Indonesia berlokasi di Pekon Bumi Ayu, Kecamatan Pringsewu, Kabupaten Pringsewu, dengan luas lahan 23.060 m² dengan keseluruhan zona buang sampah 6.210 m². Namun, proses pengelolaan sampah yang dilakukan masih cenderung dilakukan secara manual dan reaktif, di mana sampah yang dihasilkan dari tempat-tempat umum di Kabupaten Pringsewu dan terkumpul di TPS-TPS di beberapa titik di Kabupaten Pringsewu akan berakhir di TPA Bumiayu sehingga lahan sampah dikelola agar dapat menampung semua sampah yang masuk setiap hari. [3]



Gambar 1.1 Jumlah Sampah TPA Bumiayu Per Bulan (2021-2024)

Berdasarkan data historis volume sampah per bulan dari tahun 2021 hingga 2024 dapat dilihat pada gambar 1.1 terlihat adanya tren peningkatan secara signifikan. Pada awal tahun 2021 volume sampah bulanan berkisar ± 2.000 m³, kemudian mengalami peningkatan signifikan pada pertengahan 2022 hingga jumlahnya stabil di atas 3.000 m³ dan puncaknya melebihi 3.700 m³. Tren menunjukkan adanya potensi mengalami masalah terkait kapasitas lahan pengelolaan sampah di masa mendatang jika jumlah sampah terus meningkat setiap tahun dengan tidak diimbangi strategi pengelolaan yang efektif dan berbasis data.

Berdasarkan informasi dari pihak pengelola, TPA Bumiayu memiliki batas maksimal operasional jumlah sampah harian yang harus dijaga agar tidak

melampaui batas maksimal jumlah sampah harian sesuai yang telah ditentukan pada izin pengelolaan lingkungan yang berlaku. Jika jumlah sampah terus meningkat tanpa adanya antisipasi atau strategi prediktif dapat berisiko terjadi over kapasitas yang berpotensi mengganggu kelancaran operasional, penumpukan sampah, dan dampak negatif terhadap lingkungan.

Sehingga diperlukan perubahan dengan pendekatan prediktif berbasis data untuk membantu memperkirakan jumlah sampah di masa mendatang yang dapat membantu pengelola TPA dalam mengambil keputusan dengan tepat dan cepat. Dalam mendukung perubahan tersebut, salah satu langkah berbasis teknologi seperti *deep learning* dapat membantu memberikan potensi besar untuk membangun model prediksi berdasarkan pola historis dari data yang tersedia.

Di antara berbagai metode *deep learning* yang dapat digunakan untuk memodelkan data deret waktu, Long Short Term Memory dipilih karena menyesuaikan dengan kemampuannya dalam menangani pola jangka panjang. *Long Short-Term Memory* (LSTM) adalah bagian arsitektur jaringan syaraf tiruan (RNN) yang dirancang untuk mengatasi masalah vanishing gradient yang umum terjadi pada RNN konvensional sehingga lebih stabil dalam mempelajari urutan data yang panjang tanpa kehilangan konteks historis [4]. LSTM juga memiliki kemampuan dalam mempertahankan informasi penting dalam data sekuensial sehingga efektif digunakan untuk model temporal yang kompleks dan tidak linier. Berdasarkan kemampuan tersebut maka LSTM dipilih untuk memprediksi jumlah sampah yang tidak stasioner dengan memiliki pola musiman, tren jangka panjang, dan juga dapat berubah seiring waktu karena faktor eksternal[5].

Pada data jumlah sampah harian di TPA Bumiayu ini juga memiliki skala data menengah dengan jumlah data yang masih terbatas. Lebih lanjut jumlah sampah tidak hanya bergantung pada waktu tetapi juga dipengaruhi faktor eksternal seperti event harian, cuaca, jumlah penduduk, kondisi ekonomi yang mana tetap perlu dikaji kontribusinya terhadap jumlah sampah[6]. Sedangkan jika dibandingkan dengan model regresi linier atau model klasik seperti ARIMA

memiliki keterbatasan dalam menangkap informasi jangka panjang dan variabel eksternal di sisi lain[7] jika menggunakan model deep learning generatif atau seperti transformer-based membutuhkan data dalam jumlah besar dan sumber daya komputasi yang tinggi [4]. Oleh karena itu LSTM dipilih dalam penelitian ini karena mampu untuk diterapkan dengan faktor eksternal yang ada [6].

Maka penelitian ini bertujuan untuk dapat membangun dan mengevaluasi model prediksi jumlah sampah harian di TPA Bumiayu menggunakan model LSTM serta menguji pengaruh fitur eksternal dan dapat memberikan gambaran jumlah sampah di masa depan. Dengan adanya penelitian ini diharapkan dapat memberikan gambaran jumlah sampah pada TPA Bumiayu di masa depan dengan mempertimbangkan pengaruh faktor eksternal yang ada sehingga pengelola TPA dapat mengambil langkah yang preventif dan strategis. Dan juga diharapkan penelitian ini dapat berkontribusi dalam mendukung pengelolaan sampah yang lebih efisien, aktif, dan berkelanjutan serta memperkuat kesiapan daerah untuk menghadapi tantangan lingkungan di masa mendatang.

1.2 Rumusan Masalah

Adapun rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana membuat sistem kecerdasan buatan dengan model LSTM untuk mampu memprediksi jumlah sampah berdasarkan data historis sampah di TPA Bumiayu?
2. Bagaimana efektivitas dan akurasi model LSTM dalam memprediksi jumlah sampah harian berdasarkan data historis sampah di TPA Bumiayu?
3. Bagaimana pengaruh dari penambahan fitur variabel eksternal seperti data cuaca, data penduduk, dan data ekonomi terhadap akurasi model dalam memprediksi jumlah sampah harian di TPA Bumiayu?
4. Bagaimana menampilkan hasil prediksi jumlah sampah setelah mengimplementasi model LSTM selama 2025-2030 dari TPA Bumiayu?

1.3 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem kecerdasan buatan dengan model LSTM yang mampu memprediksi jumlah sampah berdasarkan data historis jumlah sampah harian di TPA Bumiayu.
2. Menganalisis performa model LSTM dalam memprediksi jumlah sampah harian berdasarkan data historis sampah di TPA Bumiayu.
3. Menganalisis model LSTM jika dilakukan penambahan fitur variabel pengaruh faktor cuaca, jumlah penduduk, dan pendapatan di Kabupaten Pringsewu dalam memprediksi jumlah sampah di TPA Bumiayu.
4. Mengembangkan dashboard visualisasi hasil prediksi jumlah sampah setelah diimplementasikan dengan model LSTM selama 2025-2030 dari TPA Bumiayu.

1.4 Manfaat Penelitian

Adapun manfaat penelitian ini adalah sebagai berikut:

1. Akademis

Mendapatkan pengalaman dalam pengembangan model LSTM dalam memprediksi jumlah sampah harian dan jangka panjang bagi peneliti. Serta memperkaya kajian akademik sebagai kontribusi terhadap literatur ilmiah dalam bidang kecerdasan buatan terutama penerapan *deep learning* dalam konteks pengelolaan lingkungan dan keberlanjutan

2. Praktis

Memberikan rekomendasi kepada pemerintah daerah, khususnya Dinas Lingkungan Hidup Kabupaten Pringsewu, dalam merancang strategi untuk meningkatkan efisiensi pengelolaan sampah berdasarkan prediksi berbasis data.

3. Teknis

Menjadi kontribusi teknis dalam membangun alur prediksi dari mulai pemrosesan data sampai dengan perancangan model LSTM serta dapat menjadi penelitian lebih lanjut pada kasus prediktif lainnya.

1.5 Batasan Masalah

Adapun batasan masalah yang menjadi fokus pada penelitian ini adalah sebagai berikut:

1. Penelitian ini mengimplementasikan model Long Short Term Memory (LSTM) dengan arsitektur Single, Double, dan Bidirectional LSTM dalam memprediksi jumlah sampah di TPA Bumiayu.
2. Dataset yang digunakan dalam penelitian ini hanya mencakup data jumlah sampah harian di TPA Bumiayu dari tahun 2021 - 2024, sedangkan data cuaca harian kabupaten Pringsewu, data jumlah penduduk kabupaten Pringsewu, dan data pendapatan perkapita kabupaten Pringsewu akan diukur setelah mendapat model terbaik.

1.6 Sistematika Penulisan

Sistematika penulisan laporan ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Memberikan penjelasan tentang latar belakang dan rumusan masalah dari memprediksi jumlah sampah LSTM di Tempat Pemrosesan Akhir, menjelaskan alasan untuk menggunakan metode LSTM, menjelaskan batasan masalah, dan menulis sistematis penulisan laporan.

BAB II : TINJAUAN PUSTAKA

Membahas dasar teori dari penelitian memprediksi jumlah sampah menggunakan LSTM, serta penelitian sebelumnya

BAB III : METODE PENELITIAN

Membuat langkah-langkah yang digunakan dalam membuat model untuk memprediksi jumlah sampah menggunakan LSTM

BAB IV : HASIL DAN PEMBAHASAN

Menjelaskan hasil dan pengujian model untuk memprediksi jumlah sampah menggunakan metode LSTM, dengan nilai yang digunakan pada metode LSTM.

BAB V : KESIMPULAN DAN SARAN

Memberikan kesimpulan dari penelitian prediksi jumlah sampah di Tempat Pemrosesan Akhir menggunakan LSTM, dan memberikan rekomendasi untuk hasil penelitian yang dapat digunakan untuk mengembangkan penelitian selanjutnya.

DAFTAR PUSTAKA

LAMPIRAN

II. TINJAUAN PUSTAKA

2.1 Sampah

Sampah adalah sisa dari proses produksi industri maupun domestik (rumah tangga). Selain itu, menurut UU Nomor 18 Tahun 2008 tentang Pengelolaan Sampah menggambarkan sampah sebagai sisa aktivitas sehari-hari manusia atau benda padat atau semi padat yang berasal dari zat organik atau anorganik yang dapat diuraikan atau tidak diuraikan dan dibuang ke lingkungan [8].

Jenis sampah menurut sifatnya

- a. Sampah organik adalah sampah yang berasal dari sisa-sisa makhluk hidup seperti manusia, tumbuhan, dan hewan yang dapat diuraikan secara alami di alam[8]. Sampah organik biasanya disebut sebagai sisa makanan, potongan sayur dan buah, sisa daun dari pepohonan, rumput-rumputan, dan sisa panen padi. Sampah organik juga dibagi menjadi sampah organik basah dan sampah organik kering. Sampah organik basah memiliki kadar air yang lebih tinggi daripada sampah organik kering, sehingga lebih mudah membusuk dan hancur.

- b. Sampah Anorganik

Sampah anorganik adalah sampah yang sulit terurai karena berasal dari bahan non-alami atau sisa limbah berbahan sintetik tertentu[8]. Contohnya adalah styrofoam, kantong plastik, aluminium, botol kaca, botol plastik, kardus, karton, tekstil, dll. Sampah dari bahan-bahan ini sulit membusuk secara alami, jadi perlu diolah kembali dengan bantuan mesin atau manusia untuk dapat digunakan kembali.

Sampah berdasarkan sumbernya[9] terbagi menjadi sebagai berikut:

- a. Sampah pemukiman pada umumnya berasal dari sampah rumah tangga seperti sisa makanan, peralatan rumah tangga, kertas, kardus, gelas pecah, kain yang tidak terpakai, sampah halaman, dan lain-lain.
- b. Sampah hasil pertanian dan perkebunan ini termasuk dalam sampah organik seperti jerami dan sebagainya.
- c. Sampah sisa material dan sisa konstruksi selama pembangunan menghasilkan sampah berbahan organik seperti kayu, bambu, triplek maupun bahan anorganik seperti semen, pasir, batu bata, ubin, besi, baja, kaca, dan kaleng.
- d. Sampah dari perdagangan dan perkantoran, sampah dari perkantoran seperti dari lembaga pendidikan, kantor pemerintah dan swasta biasanya kertas, alat tulis, pita printer, kotak tinta printer, baterai, bahan kimia laboratorium, komputer rusak, dan lain-lain. Sedangkan sampah dari perdagangan seperti dari gerai, pasar konvensional, restoran dan swalayan biasanya sampah kardus, pembungkus, kertas, sisa makanan.
- e. Sampah industri berasal dari seluruh proses produksi suatu produk, dapat berupa bahan-bahan kimia, serpihan/potongan dari pengemasan dan produk berupa kertas, kayu, plastik, kain/lap. Sedangkan untuk bahan kimia biasanya perlu cara khusus sebelum dibuang guna tidak mencemari lingkungan.

Sampah selain terbagi berdasarkan sifat dan sumbernya terdapat juga sampah khusus yang perlu penanganan khusus untuk mencegah dampak berbahaya yang dapat ditimbulkan diantaranya[9].

- a. Sampah dari rumah sakit ini seperti sisa operasi misalnya pisau bedah, botol infus dan obat-obatan harus dilakukan penanganan khusus untuk menghindari kontaminasi dengan bakteri, virus, dan bahan beracun lainnya yang dapat membahayakan manusia maupun lingkungan.
- b. Akumulator dan baterai kering ini biasanya asalnya dari rumah tangga yang terdapat bahan logam berat seperti kadmium dan raksa yang berbahaya bagi kesehatan.

2.2 TPAS Bumi Ayu

TPAS Bumi Ayu adalah tempat pemrosesan akhir sampah yang berlokasi di pekon Bumiayu kecamatan Pringsewu kabupaten Pringsewu provinsi Lampung yang telah ditetapkan sebagai TPA oleh Kabupaten Tanggamus sejak 2007. TPAS ini memiliki luas lahan keseluruhan yaitu 2,3 hektar yang jauh dari pemukiman penduduk memiliki zona pemrosesan dengan panjang tiap sisi 130 meter, 100 meter, 38 meter, dan 70 meter. TPA Bumiayu dilengkapi fasilitas seperti IPAL (Instalasi Pengolahan Air Limbah) dan IPLT (Instalasi Pengolahan Lumpur Tinja) sebagai bagian dari TPA Bumiayu. Dikelola oleh Dinas Lingkungan Hidup kabupaten Pringsewu. TPAS ini sebagai tempat pemrosesan akhir sampah-sampah yang sudah tidak dapat dikelola lagi dari TPS 3R dan TPST yang ada di Kabupaten Pringsewu.

2.3 Data Berkala (*Time Series*)

Data berkala (*time series*) atau data urutan waktu adalah data yang dikumpulkan dari satu sumber yang sama dengan interval waktu yang tidak berubah berdasarkan indeks waktu t dari waktu ke waktu dapat berupa hari, minggu, bulan, tahun, dan sebagainya. Data *time series* biasanya digunakan untuk menggambarkan suatu perubahan yang berkaitan dengan keadaan/kejadian/aktivitas. Contohnya adalah data jumlah produksi gas alam per tahun, harga emas per hari, jumlah populasi per tahun, dan masih banyak lagi. Karena pola pergerakan data dapat diketahui maka data *time series* cocok digunakan untuk pengambilan keputusan, perencanaan masa depan, dan peramalan [10].

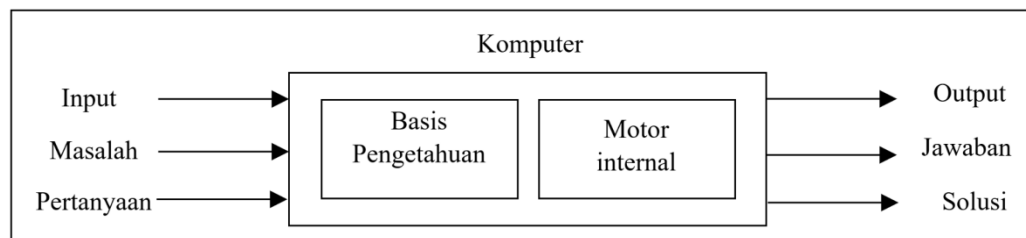
2.4 Artificial Intelligence

Artificial Intelligence dikenal juga sebagai Kecerdasan buatan adalah cabang ilmu komputer di mana mesin (komputer) memiliki kemampuan yang sama dengan dan sebanding dengan manusia untuk melakukan tugas tertentu. Meskipun awalnya diciptakan hanya sebagai alat menghitung, seiring berjalanya waktu komputer

telah berkembang menjadi alat yang dapat melakukan hampir semua tugas yang dapat dilakukan oleh manusia. Kecerdasan buatan memiliki beberapa keuntungan dibandingkan dengan kecerdasan alami manusia diantaranya:

- a. Kecerdasan buatan lebih tahan terhadap perubahan sehingga tidak cepat mengalami perubahan sepanjang sistem komputer tidak diubah lain halnya manusia yang memiliki sifat pelupa.
- b. Kecerdasan buatan lebih murah dibanding kecerdasan alami karena komputer lebih mudah digunakan dan lebih murah untuk mengerjakan suatu pekerjaan dibanding harus mendatangkan orang-orang dalam waktu tertentu.
- c. Kecerdasan buatan lebih mudah disebarkan dan di duplikat karena pengetahuan manusia membutuhkan waktu yang lama dan tidak semua pengetahuan dapat disimpan secara lengkap sedangkan pengetahuan yang disimpan dalam sistem komputer hanya perlu disalin dari komputer satu ke komputer lain dengan mudah.
- d. Kecerdasan buatan tetap bertahan sebagai bagian dari teknologi komputer sedangkan kecerdasan alami tidak dapat selalu sama.
- e. Kecerdasan buatan memiliki kemampuan menyelesaikan tugas dengan lebih baik dan lebih cepat dibanding kecerdasan alami[11].

Pada penerapan kecerdasan buatan terdapat basis pengetahuan (*knowledge base*) dan juga sebuah mesin inferensi (*inference engine*) yang jika digambarkan dapat dilihat pada Gambar 2.1



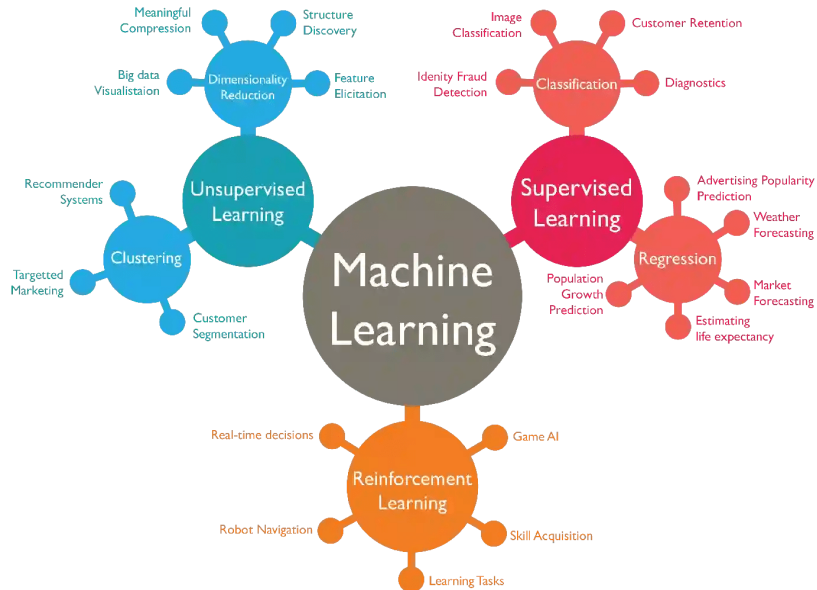
Gambar 2.1 Penerapan Konsep Kecerdasan Buatan [11]

Pada Gambar 2.1 terlihat terdapat *input* atau masalah atau pertanyaan dalam komputer di mana di dalamnya sistem komputer terdapat basis pengetahuan yang

berisi data-data, teori, maupun hubungan satu dengan yang lain dan juga terdapat motor internal yang memiliki kemampuan menarik sebuah kesimpulan berdasarkan pengalaman yang ada dan jika *input* atau masalah atau pertanyaan ini melalui sistem komputer akan menghasilkan *output* atau jawaban atau solusi dari masalah yang ada.

2.5 Machine Learning

Machine learning adalah bagian dari kecerdasan buatan yang memungkinkan komputer untuk mempelajari informasi dari data dan membuat prediksi atau tindakan berdasarkan interpretasi data tersebut. Ini memungkinkan komputer tanpa perlu pemrograman secara jelas untuk mengambil keputusan atau melakukan tindakan berdasarkan analisis data (Gregorius N. Elwirehardja; Teddy Suparyanto; Bens Pardamean, 2023). *Machine learning* ini muncul pada era big data. Mesin sederhananya "mengambil pelajaran" dari situasi sebelumnya dan menggunakan informasi ini untuk menghadapi situasi baru. Pada *machine learning* terdapat proses *training* atau pelatihan membentuk suatu model dari suatu dataset dan juga proses *testing* atau pengujian yang dilakukan setelah model dilatih untuk mengukur kinerja model [12]. Dalam *machine learning* yang terbagi dalam 3 kategori di dalamnya memiliki cakupan yang berbeda-beda seperti terlihat pada Gambar 2.2.



Gambar 2.2 Kategori *Machine Learning* [13]

Algoritma *machine learning* terbagi menjadi 3 kategori yaitu *Supervised Learning*, *Unsupervised Learning*, dan *Reinforcement Learning*[14].

- 1) *Supervised Learning* adalah algoritma *machine learning* yang setiap datanya dilabeli. Biasa untuk mengatasi masalah klasifikasi dan regresi.
- 2) *Unsupervised Learning* adalah algoritma *machine learning* pada data yang digunakan tidak membutuhkan label. Pada algoritma ini bekerja dengan menemukan pola dan informasi yang belum terbagi serta dapat menarik kesimpulan dari kumpulan data yang digunakan. Algoritma ini biasa digunakan pada *clustering* dan jaringan saraf tiruan.
- 3) *Reinforcement Learning* adalah algoritma *machine learning* yang bekerja berdasarkan interaksi dengan lingkungan dengan model yang belajar dari kesalahan dan percobaan agar dapat mencapai tujuan tertentu dengan melakukan secara optimal dari tindakan yang diambil.

2.6 Deep Learning

Deep learning merupakan salah satu bagian *artificial intelligence* yang terinspirasi dari cara kerja otak manusia. *Deep learning* ini dapat mengolah data mentah kemudian menghasilkan pola untuk pengambilan keputusan dengan

sangat efektif. *Deep learning* sebagai bagian dari *machine learning* memiliki jaringan tersendiri yang mampu memahami pola dan suatu informasi tanpa pemantauan dari data yang tidak terlabeli dan tidak terstruktur[15]. *Deep Learning* sudah banyak memberikan kemajuan dalam berbagai bidang seperti pada pemrosesan bahasa alami, pengenalan gambar, dan permainan AI. Dan *deep learning* juga algoritma yang paling efektif untuk mempelajari fitur dari data yang kompleks dan sangat besar. Contoh penerapan *deep learning* pada teknologi *self driving car* di mana mobil yang berjalan tanpa adanya pengemudi tetapi mampu mengenali rambu-rambu lalu lintas seperti halnya manusia yang mengemudi. Deep learning merupakan jaringan saraf tiruan yang deep dikatakan deep karena memiliki banyak layer deep learning semakin berkembang setelah meningkatnya teknologi seperti cpu dsb

2.6.1 Jaringan Saraf Tiruan

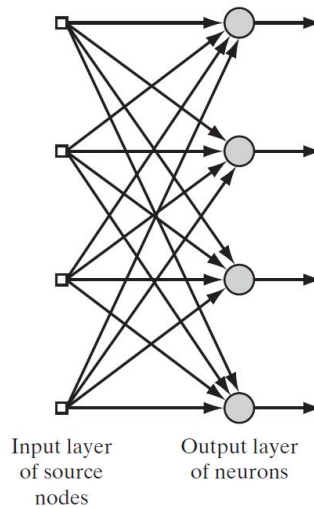
Jaringan saraf tiruan dapat digambarkan sebagai otak buatan di mana dapat berpikir seperti manusia berdasarkan potongan informasi yang diterima yang dikembangkan dalam komputer agar dapat berpikir selayaknya manusia berpikir dengan meniru aktivitas-aktivitas yang mungkin terjadi seperti mengingat, memahami, menyimpan, dan memanggil sesuatu yang telah dipelajari otak. Namun untuk sama dengan kinerja otak manusia terlalu sulit sehingga pada jaringan saraf tiruan hanya mengambil ide dari cara kerja jaringan saraf alami yang cara kerja adalah dengan mengaktifkan neuron lalu membuat koneksi, dan melatih kembali pola-pola yang ada pada otak manusia. Sedangkan untuk cara kerja jaringan saraf tiruan ini dengan neuron akan mengubah informasi atau pesan yang diterima menjadi kode. Namun jaringan saraf tiruan ini tidak di program untuk memberikan keluaran khusus jadi keluaran dihasilkan oleh pengalaman selama proses pembelajaran diikuti dan pada proses pembelajaran tersebut akan dimasukkan pola-pola *input* dan *output* kemudian jaringan akan mempelajarinya sehingga dapat memberikan jawaban yang dapat diterima [11].

Pada jaringan saraf tiruan tersusun terdapat tiga lapisan yang yaitu lapisan *input*, lapisan tersembunyi, dan lapisan *output*. Lapisan input untuk menerima informasi atau data dari luar. lapisan tersembunyi untuk memproses informasi. dan lapisan output sebagai lapisan yang menghasilkan keluaran atau JST terhadap suatu masalah. Setiap neuron dari lapisan tersembunyi dan lapisan output terhubung melalui bobot sinaptik (*weight*) yang akan mengukur seberapa besar hubungan antar neuron. Proses perhitungan dalam neuron tersebut adalah dengan menjumlahkan input yang telah dikalikan dengan bobot lalu dilanjutkan ke penerapan fungsi aktivasi untuk menghasilkan *output*. Selain itu jaringan saraf tiruan terbagi menjadi dua kelompok yang terdiri dari jaringan saraf tiruan maju (*feedforward networks*) yaitu tidak memiliki *loop* dan jaringan saraf tiruan umpan balik/berulang (*recurrent/feedback networks*) yaitu jaringan saraf tiruan yang terdapat *loop* di dalamnya[11].

Menurut Haykin, arsitektur ANN terbagi menjadi tiga yaitu jaringan umpan maju lapisan tunggal (*Single-Layer Feedforward Networks*), jaringan umpan maju berlapis-lapis (*Multilayer Feedforward Networks*), dan jaringan umpan balik (*Recurrent Networks*).

1) *Single-Layer Feedforward Network*

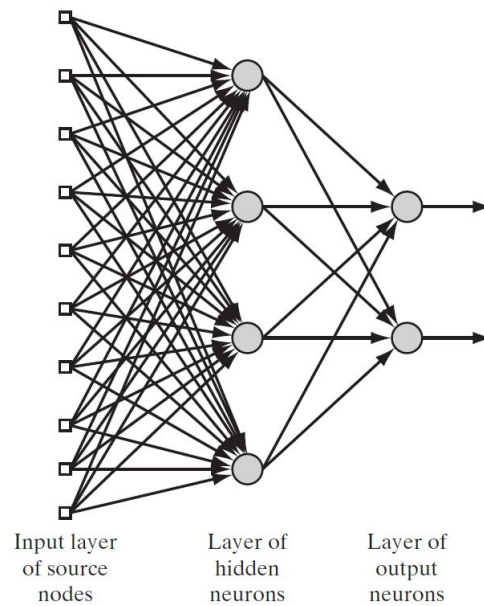
Arsitektur jaringan ini memiliki dua lapisan yaitu lapisan *input* dan lapisan *output* yang tersusun atas neuron-neuron yang membentuk lapisan, pada lapisan input akan menerima dan mengirim sinyal data dari luar sedangkan lapisan output akan memberikan hasil. Pada lapisan input tidak ada proses komputasi (proses aktivasi atau pembobotan) hanya lapisan output yang melakukan komputasi sehingga meskipun ada dua lapisan maka tetap dikatakan sebagai jaringan lapisan tunggal. Pada jaringan ini aliran sinyal dalam jaringan berlangsung satu arah dari input ke output di mana setiap neuron di lapisan output menerima sinyal dari semua neuron input dan menghasilkan keluaran berdasarkan kombinasi linier dari input dengan bobot yang kemudian dilanjutkan ke fungsi aktivasi.



Gambar 2.3 Jaringan Umpan Maju Lapis Tunggal

2) *Multilayer Feedforward Network*

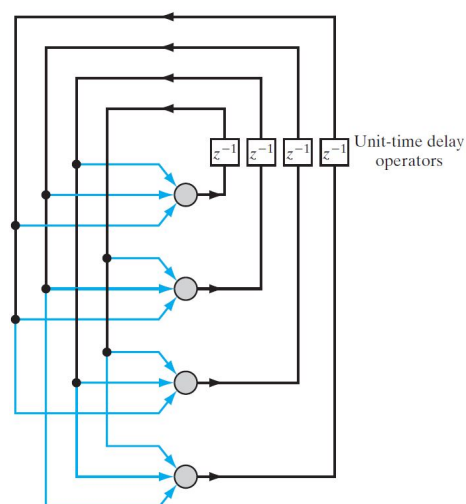
Pada arsitektur jaringan ini menambahkan satu atau lebih lapisan antara lapisan input dan lapisan output yang disebut hidden layer atau lapisan tersembunyi. Penambahan hidden layer ini dilakukan untuk mengatasi hubungan non linear atau masalah yang lebih kompleks. Pada jaringan ini aliran data dimulai dari input layer lalu diteruskan ke hidden layer pertama kemudian diteruskan ke hidden layer berikutnya jika ada lalu sampai ke output layer, dengan setiap neuron pada suatu lapisan tersembunyi menerima masukan dari lapisan sebelumnya untuk memprosesnya menggunakan bobot dan bisa lalu diterapkan kepada fungsi aktivasi sehingga menghasilkan output. Output dari lapisan tersembunyi akan menjadi input bagi lapisan selanjutnya sampai mencapai lapisan output dan menghasilkan keluaran sesuai yang diinginkan.



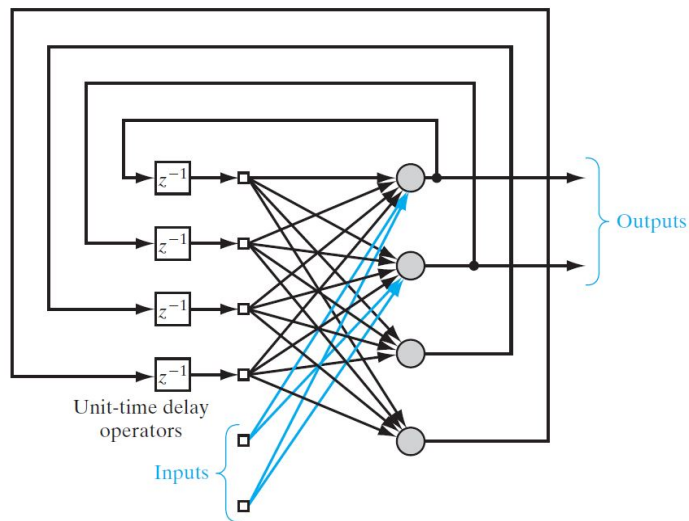
Gambar 2.4 Jaringan Umpan Maju Multilapis

3) *Recurrent Network*

Jaringan umpan balik (Recurrent Network) ini memiliki koneksi berulang (loop) dalam jaringannya. Pada arsitektur ini, output dari neuron tidak hanya diteruskan ke neuron di lapisan selanjutnya tetapi juga ke neuron di lapisan sebelumnya atau lapisan itu sendiri sehingga pada jaringan ini memiliki kemampuan untuk mengingat informasi dari waktu sebelumnya karena keluaran dari neuron akan digunakan kembali sebagai input untuk waktu berikutnya.



Gambar 2.5 Jaringan Berulang Lapis Tunggal



Gambar 2.6 Jaringan Berulang Multilapis

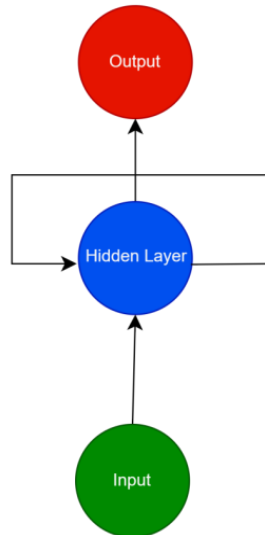
2.6.2 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) adalah jaringan saraf yang memiliki memori di mana RNN dapat mengingat hal yang terjadi di masa lalu untuk memperkirakan target dengan struktur pola waktu. Selain itu RNN juga cocok digunakan dalam masalah *Supervised* dan *Unsupervised Learning* pada data yang bersifat sekuensial seperti prediksi deret waktu. Kemampuan RNN dalam mempelajari urutan cocok digunakan pada data deret waktu yang memuat informasi yang bergerak dari waktu ke waktu yang lebih cocok diterapkan pada model RNN [15].

RNN adalah salah satu metode dalam jaringan *Neural Network* dan salah satu bagian *Supervised* dan *Unsupervised Deep Learning* yang mempelajari sampel data yang sudah ada. RNN juga dikenal dengan jaringan umpan balik karena pada jaringan *Neural Network* terdapat *loop* atau perulangan yang berguna sebagai umpan balik dalam jaringan tersebut. Pada jaringan RNN akan membuat *output* baru dengan menyesuaikan *output* jaringan pada waktu sebelumnya untuk menjadi *input*. Kemudian perulangan akan terjadi pada setiap *neuron* dan *layer* yang berguna dalam membentuk arsitektur RNN [15].

Pada pembuatan model RNN data yang digunakan bersifat sekuensial atau bertahap di mana tidak semua *input* dan *output* saling bergantung sehingga menyebabkan terjadi penumpukan tugas pada jaringan *Neural Network*. RNN dapat digunakan sebagai algoritma yang tepat dan cepat dalam mengatasi data *time series* karena RNN dapat melakukan tugas yang sama pada setiap elemen data dan memproses *output* berdasarkan hasil komputasi sebelumnya.

Pada arsitektur RNN untuk *output layer* ditandai dengan lingkaran warna merah yang akan menjadi hasil dari proses RNN dan diperoleh dari *hidden layer*. Pada *hidden layer* ditandai dengan lingkaran warna biru yang terdapat garis *looping* atau perulangan ke *hidden layer* kembali sehingga memungkinkan RNN dapat menyimpan memori yang akan dipakai pada proses selanjutnya. Dan *input layer* ditandai dengan warna hijau sebagai tempat masuknya data yang akan dipakai untuk proses komputasi RNN yang setelah data masuk maka akan ke *hidden layer*. Gambar arsitektur dapat dilihat pada gambar 2.6[15].

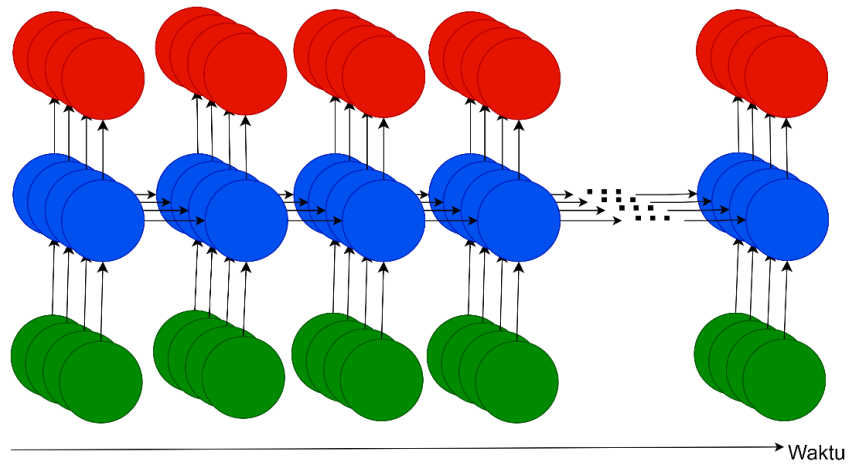


Gambar 2.7 Arsitektur Umum RNN

(Sumber: Zhou, 2019)

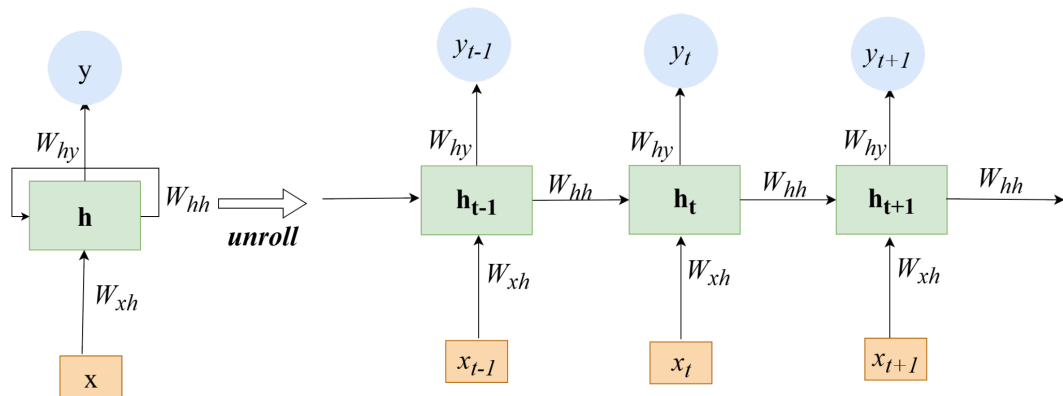
Meskipun pada gambar 2.6 arsitektur umum RNN pada *input layer* hanya terlihat memiliki satu lingkaran namun jika digambarkan lebih lengkapnya *input layer* terdiri dari banyak lingkaran (*nodes*) sedangkan seharusnya pada setiap *layer*

memiliki banyak lingkaran sedangkan gambar 2.6 dibuat untuk lebih memudahkan dalam mengilustrasikannya. Untuk melihat arsitektur RNN jika digambarkan dalam banyak *layer* dapat dilihat pada gambar 2.7.



Gambar 2.8 RNN dalam banyak *layer*

Namun karena jika RNN dilihat dalam banyak *layer* akan membingungkan dan lebih sulit digambarkan maka agar mudah dipahami arsitektur RNN juga dapat digambarkan menjadi seperti Gambar 2.8.



Gambar 2.9 Arsitektur RNN

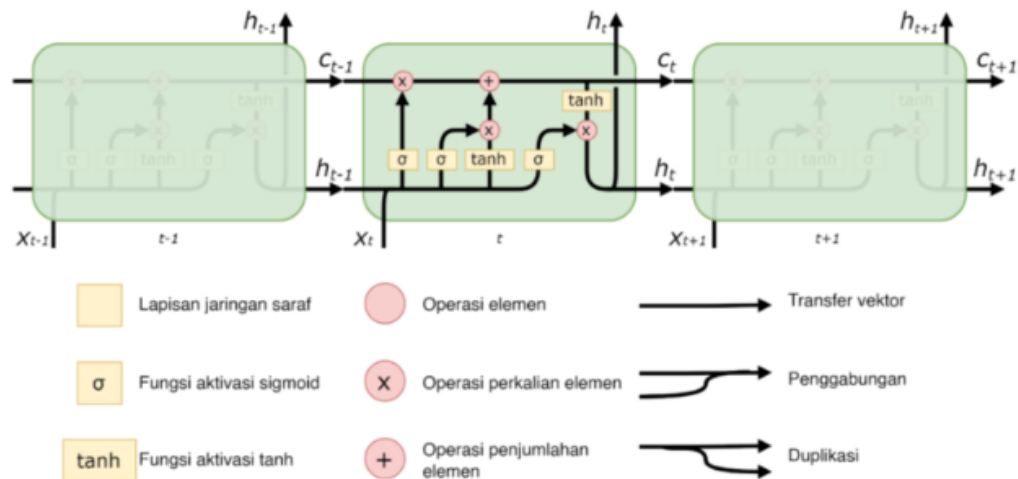
Cara kerja RNN (Recurrent Neural Network) pada gambar 2.8 dengan memproses data secara berurutan, satu per satu, sambil mempertahankan informasi dari langkah sebelumnya melalui *hidden state* (h). Setiap input (x) yang masuk akan diproses menggunakan bobot (U) untuk menghasilkan *hidden state* baru (h).

Hidden state ini juga dipengaruhi oleh *hidden state* sebelumnya menggunakan bobot (W), sehingga RNN "mengingat" informasi dari langkah sebelumnya. Selanjutnya, *hidden state* digunakan untuk menghasilkan output (O) dengan bantuan bobot (V). Jadi pada cara kerja RNN ini bekerja sama dalam memproses data dengan hubungan waktu.

2.6.3 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) yang telah diperkenalkan oleh Hohreiter dan Schmidhuber sejak tahun 1997 adalah jaringan saraf tiruan pada *deep learning* yang digunakan untuk memperbaiki masalah *backpropagation* konvensional [12]. Sebagai salah satu jenis RNN yang mampu memproses data sekuensial yang kompleks dan panjang. LSTM juga dibuat untuk mengatasi terjadinya *vanishing gradient* yang ada pada RNN yang dimodifikasi dengan menambah *memory cell* untuk menyimpan informasi jangka panjang sehingga mampu menentukan nilai *output* [14].

1. Arsitektur Long Short Term Memory



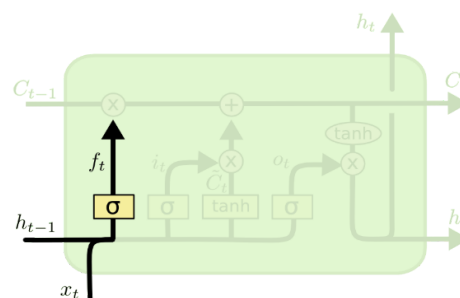
Gambar 2.10 Arsitektur LSTM [12]

Arsitektur LSTM tidak lebih sederhana dibanding arsitektur RNN sehingga LSTM memerlukan banyak komputasi karena adanya gerbang-gerbang yang mengatur proses mengingat berbagai macam hal yang terjadi pada jaringan dalam

jangka waktu panjang. Pada arsitektur LSTM kotak sebagai lapisan jaringan saraf tiruan yang di dalamnya memiliki bias dan parameter. Kemudian lingkaran sebagai operasi elemen dapat berupa perkalian atau penjumlahan elemen dengan setiap garis yang membawa seluruh vektor dari satu *node* ke *input* lain dengan vektornya dapat berupa penggabungan atau duplikasi.

Pada arsitektur LSTM juga terdapat fungsi sigmoid yang merupakan fungsi aktivasi non-linear yang mentransformasikan nilai dalam interval -1 dan 1 menjadi interval 0 dan 1 yang bermanfaat untuk melupakan atau memperbarui data jika angka yang dikali 0 maka nilainya dilupakan atau menghilang sedangkan jika dikali 1 dengan hasil sama maka nilai akan disimpan. Kemudian terdapat fungsi aktivasi Tanh sebagai fungsi aktivasi non-linear yang mengatur nilai yang melewati jaringan supaya tetap dalam interval -1 dan 1 karena vektor yang melewati jaringan akan banyak mengalami transformasi sebab berbagai operasi matematika yang dapat menghilangkan informasi maka dibutuhkan fungsi ini untuk menjadi informasi[12].

Pada LSTM terdapat beberapa gerbang atau *gate* yang terdiri dari gerbang lupa (*forget gate*), gerbang masukan (*input gate*), dan gerbang keluaran (*output gate*). Pada gerbang lupa (*forget gate*) digunakan sebagai pengontrol informasi yang perlu dilupakan atau tidak di *cell state* dengan menggunakan fungsi sigmoid sebagai fungsi aktivasi dan nilai akan diatur 1 jika informasi dipertahankan, 0 jika dilupakan, dan interval 0 dan 1 jika informasi samar [14].



Gambar 2.11 *Forget Gate*

Pada *forget gate* informasi yang dihilangkan dengan fungsi sigmoid melalui persamaan berikut.

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

Kemudian melalui persamaan di atas didapat pula persamaan:

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf})$$

Keterangan:

f_t = Forget Gate

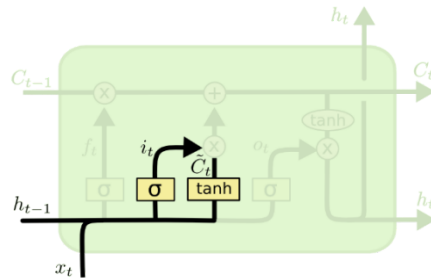
x_t = Data input

h_{t-1} = Hidden state konteks sebelumnya

W_f, W_{if}, W_{hf} = Bobot, bobot data input, bobot hidden state

b_f, b_{if}, b_{hf} = Bias, bias data input, bias hidden state

Kemudian pada gerbang masukan (*input gate*) digunakan untuk mengidentifikasi informasi penting yang akan ditambahkan pada *cell state*. Nilai dari *state* sekarang dengan *hidden state* sebelumnya akan dikirim pada fungsi sigmoid kedua selanjutnya ditransformasi antara 0 dan 1, 0 untuk informasi penting dan 1 untuk informasi yang dapat dibuang.



Gambar 2.12 *Input Gate*

Pada *input gate* didapat setelah informasi diolah pada *forget gate* lalu informasi baru didapat untuk *cell state* berikut persamaannya.

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

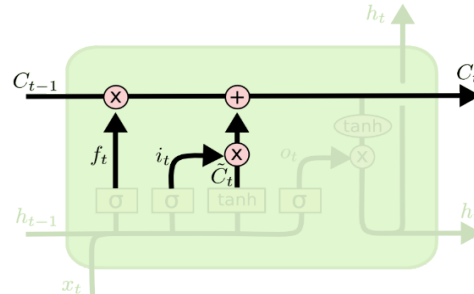
Lalu didapat persamaan

$$i_t = \sigma(W_{ii}x_t + b_{if} + W_{hi}h_{t-1} + b_{hi})$$

$$\hat{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c$$

Keterangan:

$i_t = \text{Input Gate}$

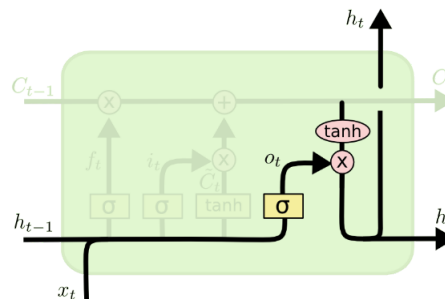


Gambar 2.13 Cell State

Persamaan pada *cell state*

$$C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t$$

Pada *output gate* digunakan untuk menentukan nilai untuk *hidden state* pada *layer* LSTM berikutnya.



Gambar 2.14 Output Cell

Output Cell dihitung dengan persamaan berikut.

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

$$h_t = o_t \times \tanh(C_t)$$

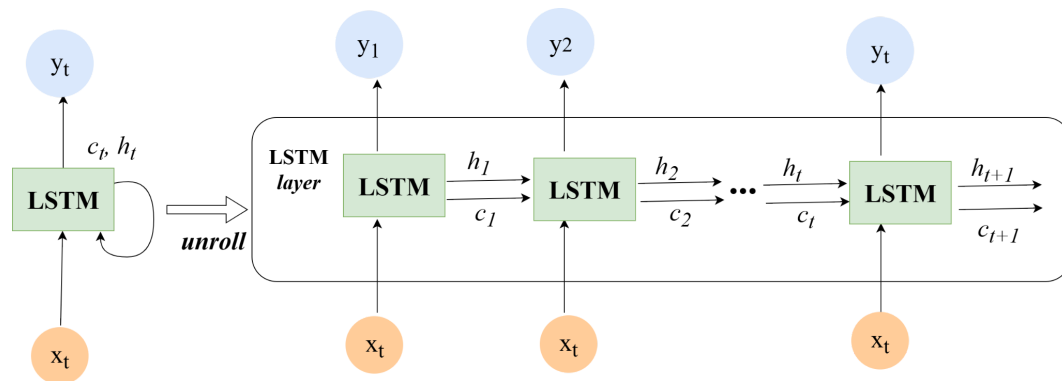
Keterangan

$o_t = \text{Output Gate}$

2. Single Long Short Term Memory

Arsitektur dasar LSTM hanya memiliki satu lapisan tersembunyi berupa unit-unit LSTM yang akan memproses data secara berurutan sehingga dapat dikatakan

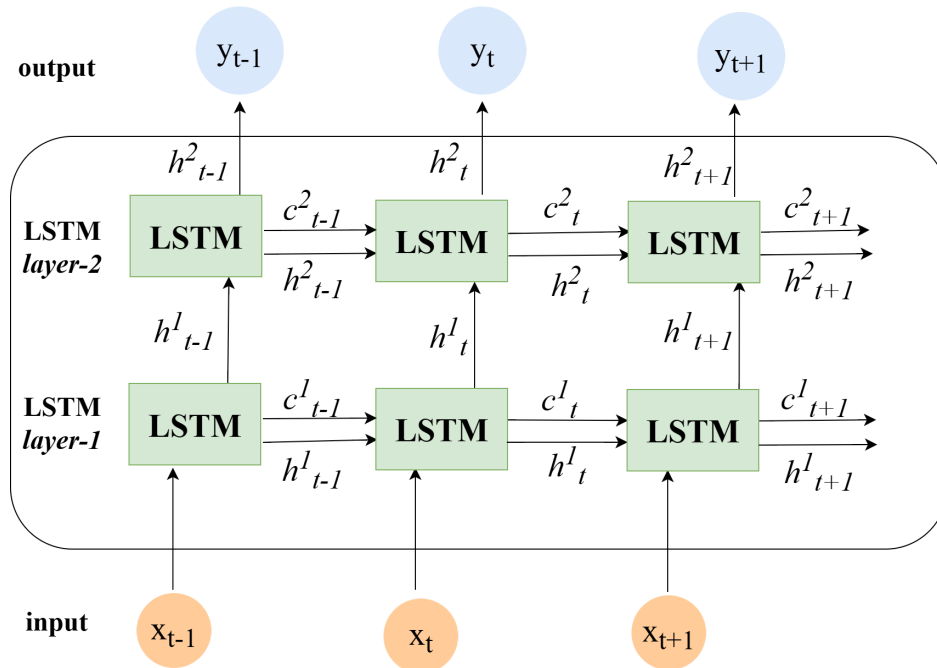
sebagai Single layer LSTM. Pada arsitektur ini input akan dimasukkan ke dalam lapisan LSTM kemudian akan diproses untuk menghasilkan output.



Gambar 2.15 *Single Layer LSTM*

3. Stacked Long Short Term Memory

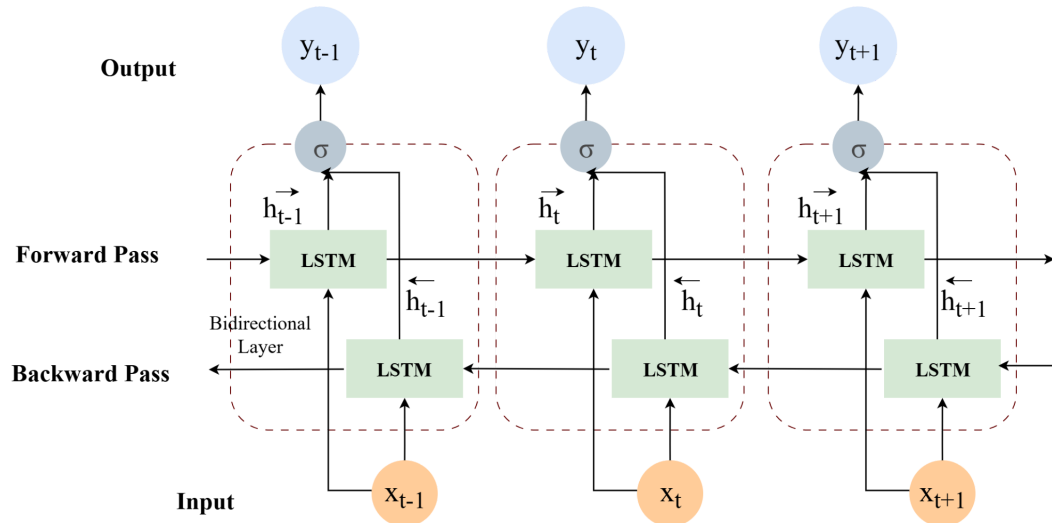
Untuk double atau stacked LSTM ini dapat dikatakan juga sebagai multilayer LSTM karena memiliki lebih dari satu lapisan LSTM tersembunyi. Pada arsitektur ini terdapat dua lapisan LSTM tersembunyi yang tersusun secara bertingkat (stack) dapat juga dua atau lebih sehingga output dari LSTM lapisan pertama akan menjadi input untuk LSTM lapisan kedua. Arsitektur ini digunakan untuk mengatasi data dengan pola yang lebih kompleks[5]. Berdasarkan arsitekturnya setelah input diproses di lapisan LSTM pertama maka outputnya akan menghasilkan *hidden state* dan memori baru yang akan diproses ke lapisan selanjutnya dan berlanjut sehingga didapatkan nilai output dari *hidden layer* atau lapisan LSTM kedua.



Gambar 2.16 Double Layer LSTM

4. Bidirectional Long Short Term Memory

Pada bidirectional LSTM ini terdapat dua jaringan LSTM yang melakukan pemrosesan data dari dua arah (maju dan mundur). Jaringan LSTM satu memproses urutan data ke arah depan (maju) atau dari awal ke akhir disebut *forward* sedangkan jaringan LSTM dua akan memproses urutan data ke arah sebaliknya atau ke arah belakang (mundur) atau dari akhir ke awal disebut *backward* [5]. Lalu kedua output dari jaringan LSTM *forward* dan *backward* itu digabung pada setiap urutan waktu. Sehingga dengan adanya dua lapisan yang memproses data secara berlawanan itu model dapat mempelajari informasi dari masa mendatang maupun masa lalu untuk setiap urutan input.



Gambar 2.17 Bidirectional LSTM

Dalam arsitekturnya terdapat untuk proses *backward* ditandai dengan panah yang mengarah ke belakang atau kiri sedangkan pada *forward* ditandai dengan panah yang mengarah ke depan atau kanan. Sehingga ketika diproses ke lapisan LSTM *forward*, LSTM akan memproses input berdasarkan urutan waktu pertama dan *hidden state* serta memori sebelumnya untuk menghitung *hidden state* dan memperbarui *cell state* begitu pula dengan pada proses *backward* tetapi data akan diproses dari paling belakang kemudian output dari masing masih proses maju dan mundur akan digabungkan untuk dimasukkan ke fungsi aktivasi agar menghasilkan sebuah keluaran.

2.7 Fungsi Aktivasi

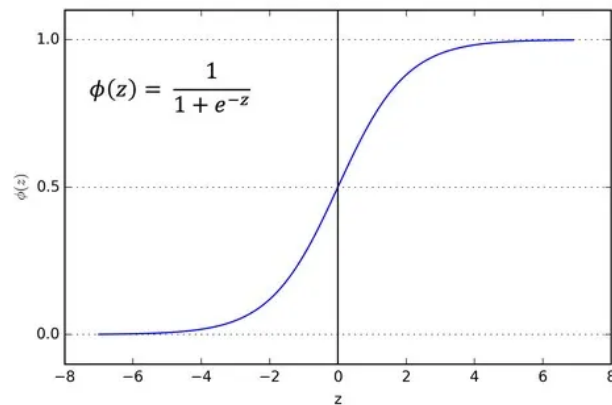
Fungsi aktivasi menentukan output suatu neuron dan juga digunakan oleh setiap neuron pada jaringan saraf tiruan untuk menentukan sinyal (nilai input) harus diteruskan ke neuron berikutnya atau tidak. Fungsi aktivasi juga menentukan suatu neuron apakah harus diaktifkan atau tidak. Fungsi aktivasi terbagi menjadi dua kategori.

1) Fungsi Aktivasi Linier

Fungsi ini memiliki bentuk $f(x) = x$ sehingga outputnya akan bernilai sama dengan nilai input. Fungsi ini nilainya rentang $-\infty$ sampai ∞ sehingga tidak dapat mampu mengatasi masalah yang kompleks.

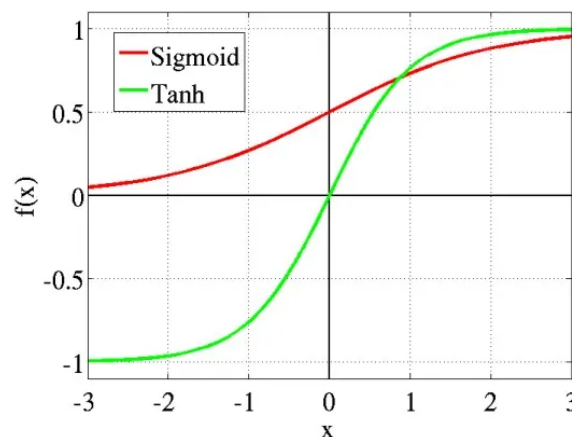
2) Fungsi Aktivasi Non-Linier

- a. Fungsi sigmoid adalah fungsi dengan rentang antara 0 sampai 1 sesuai digunakan untuk memprediksi probabilitas.



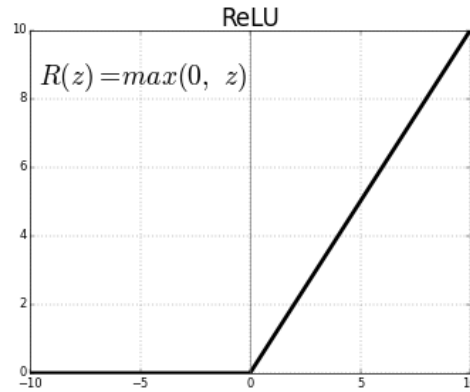
Gambar 2.18 Kurva Fungsi Sigmoid

- b. Fungsi Tanh atau Hyperbolic Tangent ini membentuk S dengan rentang antara -1 sampai 1. Fungsi ini memetakan masukan nol ke nilai mendekati 0 dan nilai negatif ke masukan yang sangat negatif.



Gambar 2.19 Kurva Fungsi Tanh

- c. Fungsi ReLU atau Rectified Linear Unit paling banyak digunakan. Keluaran pada fungsi ini linear jika input positif dan nol jika input negatif sehingga nilainya 0 sampai ∞ .



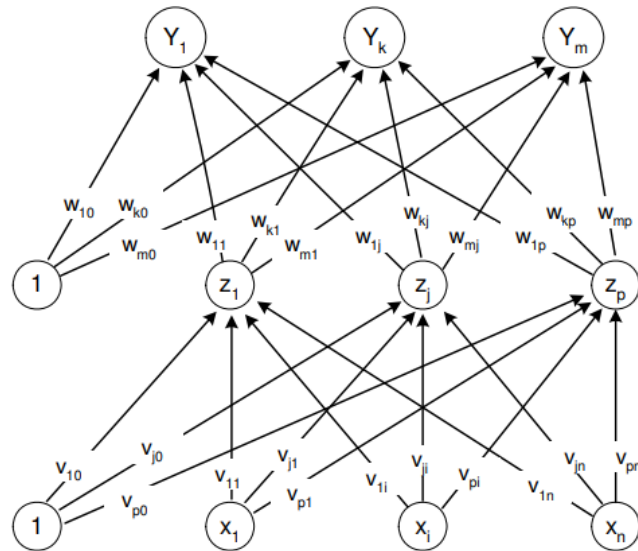
Gambar 2.20 Kurva Fungsi ReLU

2.8 Backpropagation

Jaringan saraf tiruan satu lapisan memiliki kesulitan dalam mengenali pola. Namun kelemahan ini dapat diperbaiki dengan menambahkan satu atau lebih lapisan tersembunyi antara lapisan input dan lapisan output. Meskipun biasanya mulai mencoba dengan satu layar tersembunyi terlebih dahulu, penggunaan lebih dari satu layar tersembunyi memiliki keuntungan dalam beberapa kasus karena memerlukan waktu yang lama dalam pelatihan. Seperti model JST lainnya, backpropagation mengajarkan jaringan untuk mengimbangi kemampuan dalam mengenali pola yang digunakan selama pelatihan dan memberikan respons yang tepat terhadap pola masukan yang serupa, tetapi tidak sama dengan pola yang digunakan selama pelatihan. [16]

Backpropagation adalah algoritma pembelajaran jaringan saraf tiruan. Untuk mencapai target yang diinginkan bobot-bobot jaringan saraf tiruan dengan arah mundur disesuaikan berdasarkan nilai kesalahan (error) yang ditemukan dalam proses pembelajaran. Backpropagation membutuhkan tiga lapisan (layer) utama yaitu lapisan masukan (input layer) menghubungkan jaringan ke sumber data luar, lapisan tersembunyi (hidden layer) memiliki lebih dari satu lapisan tersembunyi atau bahkan tidak sama sekali, dan lapisan luaran (output layer), mengambil hasil

dari lapisan input melalui fungsi sigmoid. Output dari lapisan ini yang dianggap hasil dari proses. Pembelajaran terawasi juga dikenal sebagai backpropagation dengan membuat fungsi dari data pelatihan untuk mempelajari fungsi pemetaan dari input ke output. Dengan kata lain, pembelajaran terawasi memiliki target (Goal) untuk memperkirakan fungsi pemetaan, sehingga algoritma dapat memprediksi output untuk input baru ketika diberikan.[17]



Gambar 2.21 Arsitektur Backpropagation

Pelatihan backpropagation terdiri dari tiga fase. Pertama, pola masukan dihitung maju dari layar masukan ke layar keluaran dengan menggunakan fungsi aktivasi yang ditentukan. Fase kedua adalah fase maju, dan fase ketiga adalah fase mundur. Kesalahan dipropagasikan mundur, dimulai dari garis di layar keluaran yang berhubungan langsung dengan unit, ketika keluaran jaringan berbeda dengan target yang diinginkan. Fase ketiga adalah mengubah bobot sehingga kesalahan berkurang.

a. Fase I : Propagasi maju

selama propagasi maju, sinyal masukan ($= x_i$) dikirim ke lapisan tersembunyi melalui fungsi aktivasi yang ditentukan. Keluaran setiap unit layar tersembunyi ($= z_j$) dikirim kembali ke lapisan tersembunyi di atasnya melalui fungsi aktivasi yang ditentukan sampai keluaran jaringan ($= y_k$) dihasilkan.

Kemudian, keluaran jaringan ($= y_k$) dievaluasi bersama dengan target yang harus dicapai ($= t_k$). Selisih antara t_k dan y_k adalah kesalahan. iterasi dihentikan jika kesalahan kurang dari batas toleransi,. Namun, bobot setiap garis jaringan akan diubah untuk mengurangi kesalahan jika kesalahan melebihi dari batas toleransi,

b. Fase II : Propagasi mundur

Berdasarkan kesalahan $t_k - y_k$, faktor Δk ($k = 1, 2, \dots, m$) dihitung dan digunakan untuk mendistribusikan kesalahan di unit y_k ke semua unit tersembunyi yang terhubung langsung dengan y_k . Selain itu, faktor Δj dihitung untuk setiap unit di lapisan tersembunyi sebagai dasar untuk mengubah bobot semua garis yang berasal dari unit tersembunyi di bawahnya.

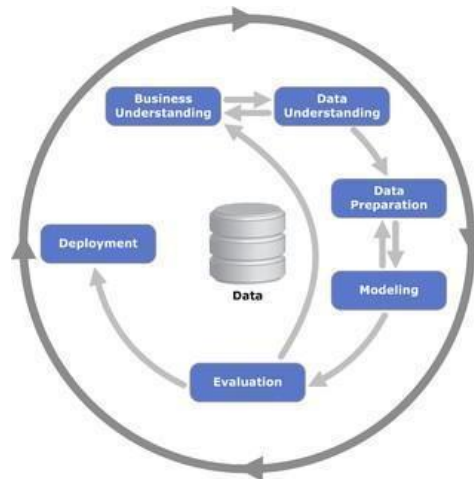
c. Fase III : Perubahan bobot

Setelah faktor delta dihitung, semua garis diberi bobot secara bersamaan. Perubahan bobot suatu garis tergantung pada faktor delta neuron di layar atasnya. Perubahan bobot garis yang menuju ke layar keluaran, misalnya, didasarkan pada faktor delta k yang ada di unit keluaran.

Ketiga fase tersebut diulang-ulang sampai kondisi penghentian dipenuhi. Jumlah iterasi atau kesalahan adalah kondisi penghentian yang paling umum. Iterasi akan dihentikan jika jumlah iterasi yang dilakukan melebihi jumlah maksimum iterasi yang ditetapkan atau jika kesalahan yang dilakukan sudah lebih kecil dari batas toleransi yang diizinkan.

2.9 CRISP-DM

Cross - Industry Standard Process for Data Mining (CRISP-DM) adalah suatu standar proses dalam data mining sebagai solusi dalam menyelesaikan masalah umum dari suatu penelitian atau bisnis. Pada CRISP-DM, proyek data mining memiliki 6 tahap atau beruntun yang sifatnya adaptif. Di mana keluaran pada fase sebelumnya akan mempengaruhi fase berikutnya [10].



Gambar 2.22 Tahapan CRISP-DM [14]

Proses metode CRISP-DM seperti yang terlihat pada Gambar 2.15 berupa siklus yang memperlihatkan urutan fase yang dapat berubah dengan adanya kemungkinan kembali ke fase sebelumnya atau pergerakan maju mundur dan juga tidak searah yang mungkin dapat terjadi pada beberapa fase [14].

Adapun penjelasan setiap fase atau tahapan pada CRISP-DM sebagai berikut [10].

1. *Business Understanding*

Fase ini merupakan tahapan cukup penting untuk menentukan tujuan proyek serta kebutuhan dengan detail, mengatur tujuan dan menentukan batasan dalam masalah data mining, dan menyusun strategi awal untuk mencapai tujuan.

2. *Data Understanding*

Fase ini dilakukan untuk menjadi tonggak pada penelitian dalam menyusun ringkasan dan melakukan identifikasi potensi terhadap masalah yang muncul pada data yang harus dilakukan secara teliti. Pada fase ini akan dilakukan pengumpulan data untuk kebutuhan proyek, melakukan analisis hasil untuk memahami lebih lanjut terkait data serta sebagai proses untuk mencari informasi awal, serta melakukan evaluasi kualitas data yang telah dihasilkan sebelumnya. Dan jika diperlukan maka dapat memilih sebagai kumpulan data yang terdapat kemungkinan dan terdapat pola masalah.

3. *Data Preparation*

Fase ini dilakukan untuk menghasilkan data yang siap untuk fase *modelling*. Pada fase ini dilakukan persiapan data dari awal untuk fase berikutnya, memilih variabel yang sesuai, mengubah variabel yang dibutuhkan, dan menyiapkan data awal hingga siap untuk pemodelan. Pada tahap ini sering dilakukan iterasi untuk menemukan hal yang sesuai dengan data yang dimaksud.

4. *Modelling*

Fase ini dilakukan penentuan teknik, alat bantu, dan algoritma data mining yang sesuai kemudian menerapkan teknik dan algoritma dengan alat bantu yang digunakan. Jika perlu menyesuaikan antara data dengan metode yang digunakan maka dapat kembali pada fase sebelumnya.

5. *Evaluation*

Fase ini merupakan fase evaluasi dengan melakukan evaluasi terhadap satu model atau lebih dari tahap pemodelan sehingga dapat menghasilkan kualitas dan efektivitas yang optimal sebelum disebarkan ke fase selanjutnya. Fase ini juga menetapkan model yang sudah mencapai tujuan pada awal fase.

6. *Deployment*

Fase ini merupakan fase penting dalam proses CRISP-DM sebagai rencana penggunaan model yang sudah dimulai sejak fase pertama dilakukan untuk menyajikan hasil yang dapat dipahami pengguna.

2.10 Evaluasi Model

Setelah melakukan pemodelan akan dilakukan evaluasi metode untuk mengukur kemampuan dari suatu model yang dibangun. Di mana evaluasi ini akan mengetahui hasil akurasi dari model prediksi [13].

a. *Root Mean Square Error (RMSE)*

RMSE adalah standar dalam pengukuran *error* yang banyak ditemukan dalam evaluasi model hasil prediksi dengan didapat dari rata-rata pangkat perbandingan antara nilai aktual dan prediksi. Dan semakin kecil nilai RMSE semakin akurat pula hasil prediksi yang dilakukan.

Persamaan *Root Mean Square Error* (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - A_i)^2}$$

Keterangan:

P_i = Data Aktual

A_i = Data prediksi

n = Jumlah Data

b. *Mean Absolute Percentage Error* (MAPE)

MAPE adalah rata-rata nilai absolut kesalahan yang muncul dari hasil prediksi berupa persentase. Untuk nilai MAPE semakin kecil nilai MAPE semakin besar kemungkinan akurasi model prediksi.

Persamaan *Mean Absolute Percentage Error* (MAPE):

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{P_i - A_i}{P_i} \right|$$

Keterangan:

P_i = Data Aktual

A_i = Data prediksi

n = Jumlah Data

Menurut Lewis (1982) mengusulkan interpretasi umum terhadap nilai MAPE dapat dilihat dalam tabel 2.1.

Tabel 2.1 Tingkat Akurasi Berdasarkan Nilai MAPE

Nilai MAPE (%)	Interpretasi
<10%	Akurasi sangat baik
10-20%	Akurasi baik
20-50%	Akurasi layak/cukup
>50%	Akurasi buruk

c. *Mean Absolute Error* (MAE)

MAE adalah rata-rata dari selisih nilai dari peramalan atau prediksi dengan data aktual dalam bentuk absolut. Nilai rentang MAE antara 0 sampai tak hingga. Semakin kecil nilai MAE semakin baik kinerja model untuk prediksi.

Persamaan *Mean Square Error* (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - A_i|$$

Keterangan:

P_i = Data Aktual

A_i = Data prediksi

n = Jumlah Data

2.11 Dekomposisi

Dekomposisi adalah metode dalam analisis data deret waktu yang memisahkan data time series untuk memberikan informasi mendalam pada setiap komponen yang dipisahkan menjadi komponen tren, musiman, siklis, dan error. Komponen tren menggambarkan perkembangan jangka panjang pada suatu periode sehingga dapat melihat pertumbuhan datanya yang didapat dari nilai rata-rata data historis dalam periode waktu tertentu. Kemudian komponen musiman (*seasonal*) ini mewakili adanya fluktuasi atau pergerakan musiman yang berulang pada data historis dengan periode yang sama seperti pola harian, bulanan, tahunan, atau kuartalan. Dan juga komponen error (*residual*) yang mewakili nilai error atau residual yang tidak dapat ditangkap oleh komponen musiman atau tren.[18]

2.12 Korelasi Pearson

Analisis korelasi adalah teknik statistik yang sangat penting untuk menilai hubungan antara dua atau lebih variabel. Menurut Freedman et al. (2007), koefisien korelasi Pearson biasanya digunakan untuk mengukur korelasi sehingga dapat menunjukkan kekuatan dan arah hubungan antar variabel. Nilai koefisien

korelasi berkisar antara -1 dan 1, dengan hubungan yang searah ditunjukkan oleh nilai positif sedangkan hubungan yang terbalik ditunjukkan oleh nilai negatif, dan hubungan yang lemah atau tidak berkorelasi ditunjukkan oleh nilai yang hampir atau sama dengan 0. Koefisien korelasi Pearson, yang digambarkan sebagai skala antara variasi berkorelasi dari variabel-variabel dan hasil perkalian dari masing-masing standar deviasi dan biasa digunakan untuk menghitung korelasi antar variabel.

Metode ini bergantung pada standar deviasi setiap variabel dan kovariansi antara keduanya. Koefisien ini memiliki nilai mulai dari -1 hingga 1, dengan nilai -1 menunjukkan hubungan linear negatif sempurna, nilai 1 menunjukkan hubungan linear positif sempurna, dan nilai 0 menunjukkan bahwa dua variabel tidak ada hubungan linear. Tidak ada korelasi yang menyebabkan kausalitas antara dua variabel. Dengan kata lain, tidak ada korelasi yang menunjukkan bahwa salah satu variabel mempengaruhi variabel lainnya. Sebaliknya, korelasi mengukur bagaimana dua variabel bergerak bersama-sama dalam pola atau tren tertentu, tetapi tidak memberikan alasan mengapa korelasi terjadi. Biasanya korelasi disajikan dalam bentuk visualisasi data seperti *heatmaps* yang digunakan untuk menunjukkan pola atau tren dalam data temporal atau spasial.[19]

2.13 Interquartile Range (IQR)

Interquartile Range (IQR) adalah salah satu ukuran penyebaran yang dapat memberikan informasi tambahan terkait keragaman data yang dimiliki. IQR adalah perbedaan antara kuartil ketiga dan kedua. Menurut Devore, kuartil adalah nilai-nilai yang membagi dataset menjadi empat bagian yang sama besar(2019). Kuartil pertama, atau Q1, adalah nilai yang membagi 25% data terendah, Kuartil kedua, atau Q2, adalah median dataset, dan Kuartil ketiga, atau Q3, adalah nilai yang membagi 25% data tertinggi. penggunaan kuartil sebagai ukuran penyebaran data memiliki keuntungan utama karena menunjukkan sebaran nilai-nilai tengah dataset, di mana sebagian besar nilai. IQR sering digunakan dalam analisis untuk mengatasi dampak outlier karena lebih tahan terhadap outlier. IQR juga tidak

terpengaruh oleh nilai ekstrim dapat digunakan untuk mengukur sebaran data di sekitar median. Jika IQR relatif sempit, itu menunjukkan bahwa sebagian besar data berada di sekitar median, dan jika IQR lebar, itu menunjukkan bahwa data tersebar secara luas di sepanjang rentang nilai. [19]

2.14 Microsoft Excel

Microsoft Excel adalah aplikasi *spreadsheet* yang dikembangkan dan didistribusikan oleh Microsoft. Ini memungkinkan pengguna mengorganisir, memformat, dan menghitung data menggunakan sistem *spreadsheet*, serta melakukan analisis dan visualisasi data dalam bentuk grafik dan tabel.

Microsoft Excel menyediakan berbagai fitur, termasuk:

- a. *Spreadsheet* dan Tabel: Membuat dan mengatur data dalam format baris dan kolom.
- b. Fungsi dan Rumus: Excel memiliki ratusan fungsi dan rumus untuk melakukan perhitungan matematis, statistik, logika, dan keuangan.
- c. Grafik dan Diagram: Pengguna dapat membuat berbagai jenis grafik untuk visualisasi data, seperti grafik batang, garis, pai, dan lainnya.
- d. PivotTable dan PivotChart: Alat canggih untuk analisis data yang memungkinkan pengguna untuk merangkum, menganalisis, mengeksplorasi, dan menyajikan data.
- e. Macro dan VBA (*Visual Basic for Applications*): Mengotomatisasi tugas berulang dengan menggunakan macro dan VBA.[20]

2.15 Google Colaboratory

Google Colaboratory, juga disebut Colab, adalah layanan Jupyter Notebook yang dihosting tanpa konfigurasi dan layanan *cloud* gratis yang memungkinkan pengguna untuk menulis, menjalankan, dan membagikan kode Python di lingkungan berbasis browser, menggabungkan kode Python dengan teks, bagan, gambar, HTML, dan LaTeX yang kaya ke dalam satu dokumen, menyimpan dokumen di Google Drive, mengakses buku catatan dari mana saja secara gratis,

dan menggunakan sumber daya komputasi, termasuk GPU dan TPU, secara gratis. Google Colab cocok digunakan untuk pendidikan, *machine learning*, dan *data science* meskipun tersedia secara gratis tetapi jika pengguna membutuhkan komputasi yang lebih maka ada pilihan untuk membayar ketika kebutuhan komputasi meningkat [21].

2.16 Streamlit

Streamlit adalah *framework open-source* berbasis Python yang memudahkan pengguna untuk membangun aplikasi berbasis web interaktif yang semuanya menggunakan python murni tanpa perlu pengalaman front-end sehingga dapat dibagikan dalam hitungan menit. Streamlit biasanya digunakan untuk keperluan data scientist dan AI/ML engineers banyak sekali hal-hal yang dapat dibangun seperti LLM dan aplikasi chatbot, Aplikasi sains dan teknologi, aplikasi NLP dan bahasa, aplikasi keuangan dan bisnis, aplikasi geografi dan masyarakat dan masih banyak lagi. Streamlit ini memiliki kelebihan yaitu sederhana dan pythonic, dapat dijadikan prototipe dengan cepat dan interaktif sehingga dapat diakses oleh orang lain, berbasis open-source dan gratis. Pada streamlit dapat mempercantik aplikasi dengan banyak elemen tambahan yang disediakan. [22]

2.17 Python

Python adalah bahasa skrip yang mudah dipelajari namun kuat dan serbaguna, yang membuatnya menarik untuk pengembangan aplikasi. Selain itu, Python adalah bahasa pemrograman dinamis dan tingkat tinggi yang berfungsi sebagai interpreter, memungkinkan program untuk mengubah kode sumber langsung menjadi kode mesin. Bahasa ini juga memiliki banyak struktur data tingkat tinggi dan mendukung pendekatan pemrograman berorientasi objek untuk pengembangan aplikasi. Karena sifat interpretasinya yang sangat dinamis, sintaks dan pengetikan Python adalah pilihan yang ideal untuk pengembangan aplikasi dan skrip yang cepat [23]. Guido van Rossum merancang bahasa ini pada tahun 1991. Python Software Foundation terus mengembangkan Python hingga saat ini. Python mendukung hampir semua sistem operasi, bahkan Linux.

2.18 Penelitian Terkait

Adapun penelitian terkait yang digunakan untuk menunjang penelitian yang dilakukan ini dapat dilihat dalam tabel 2.2.

Tabel 2.2 Penelitian Terkait

No	Deskripsi
1.	Judul: <i>Revolutionising Waste Management with the Impact of Long Short Term Memory Networks on Recycling Rate Predictions</i>
	Penulis: Matthew A. Olawumi, Bankole I. Oladapo, Rukayat Abisola Olawale
	Tahun: 2024
	Data: Data historis tingkat daur ulang dari berbagai sumber publik termasuk faktor populasi, ekonomi, kebijakan, dan musiman
	Metode: LSTM dengan dua lapisan (100 dan 50 neuron) <i>dropout</i> 200%, dilatih selama 100 epoch
	Hasil: Untuk melakukan prediksi tingkat daur ulang dalam manajemen limbah. Pada penelitian ini menggunakan beberapa variabel selain tingkat daur ulang, terdapat data demografi, faktor musiman, indikator ekonomi, dan perubahan kebijakan yang mungkin terjadi dengan menggunakan metode LSTM dan pembandingnya yaitu ARIMA dan regresi linier. LSTM lebih unggul dengan nilai MAE 3,5%, RMSE 2,8%, dan R^2 0,92.[24]
2.	Judul: <i>Multi-site Household Waste Generation Forecasting Using a Deep Learning Approach</i>
	Penulis: Maximiliano Cubillos

No	Deskripsi
	Tahun: 2020
	Data: Data mingguan berat sampah dari 1.000 rumah tangga di Herning Denmark (2011-2018) ditambah variabel eksternal: cuaca, kelembaban, tekanan udara, sinar matahari.
	Metode: Multi-site LSTM, ARIMA, dan MLP.
	Hasil: Menghasilkan model <i>multi-site</i> LSTM atau model yang melatih satu model untuk semua rumah tangga dibanding model ARIMA dengan mampu mengurangi rata-rata error sampai 85% dan akurasi lebih tinggi 28% dibanding model <i>single-site</i> , <i>multi-site</i> LSTM sendiri adalah LSTM yang menggunakan banyak variabel sedangkan <i>single-site</i> LSTM adalah LSTM yang hanya menggunakan data mingguan berat sampah dari 1.000 rumah tangga di Herning.[6]
3.	Judul: <i>Deep Learning Based: A Comprehensive Analysis for Waste Prediction</i>
	Penulis: Anil Utku dan Sema Kayapinar Kaya
	Tahun: 20222
	Data: Data harian produksi sampah rumah tangga di Istanbul (2016-2021)
	Metode: LSTM, kNN, Random Forest, SVM, MLP, GRU
	Hasil: Penelitian ini menggunakan metode LSTM untuk memprediksi jumlah sampah harian di kota Istanbul, Turki menggunakan data yang berisi tanggal dan jumlah sampah dari tahun 2016 – 2021. Kemudian model LSTM dibandingkan dengan model lain seperti kNN, RF, SVM, MLP, dan GRU dan hasilnya LSTM lebih unggul dibanding metode

No	Deskripsi
	pembandingan dengan nilai MAE atau tingkat kesalahan sebesar 5,19% lebih rendah.[25]
4.	Judul: <i>Forecasting the Status of Municipal Waste in Smart Bins Using Deep Learning</i>
	Penulis: Sabbir Ahmed, Sameera Mubarak, Jia Tina Du, Santoso Wibowo
	Tahun: 2022
	Data: Data sensor dari smart bin di lingkungan kota yang diambil dari IoT
	Metode: 1D-CNN, LSTM, GRU, Bi-LSTM
	Hasil: Penelitian ini untuk memprediksi status pengisian dengan data dari sensor <i>smart bin</i> di Wyndham, Australia dengan menggunakan metode 1D CNN, LSTM, GRU, dan Bi-LSTM di mana pada datanya mencakup waktu dan tingkat pengisian <i>smart bin</i> . Hasil penelitian menunjukkan performa terbaik pada LSTM dengan MAE, 0,602, RMSE 1,579, MAPE 1,855, dan R^2 0,925 sehingga LSTM cocok untuk memproses data sekuensial yang dapat melakukan prediksi dengan baik.[6]
5.	Judul: LSTM-Based Forecasting for Urban Construction Waste Generation
	Penulis: Li Huang, Ting Cai, Ya Zhu, Yuliang Zhu, Wei Wang, Kehua Sun
	Tahun: 2020
	Data: Data tahunan sampah konstruksi dari Shanghai dan Hong Kong
	Metode: LSTM, Ridge Regression (RR), Support Vector Regression

No	Deskripsi
	(SVR), dan Backpropagation Neural Network (BPNN).
	Hasil: Hasil penelitian ini menunjukkan bahwa model berbasis LSTM dapat menyelesaikan masalah peramalan non linier univariat dengan sangat baik. Selain itu, lapisan <i>dropout</i> ditambahkan untuk mengatasi masalah <i>overfitting</i> yang efektif dari model berbasis LSTM dan meningkatkan kinerja generalisasi. Dengan hasil dari set latihan, MAE adalah 5,54, MAPE adalah 2,83, dan RMSE adalah 7,07; untuk set uji LSTM, MAE adalah 5,48, MAPE adalah 3,98, dan RMSE adalah 5,95.[26]
6.	Judul: <i>The Use of a Recurrent Neural Network Model with Separated Time-Series and Lagged Daily Inputs for Waste Disposal Rates Modeling during COVID-19</i>
	Penulis: Hoang Lan Vu, Kelvin Ng, Amy Richter, Golam Kabir
	Tahun: 2021
	Data: Data harian disposal sampah kota Regina, Kanada (2013-2020),
	Metode: RNN
	Hasil: Model LSTM yang digunakan untuk memprediksi tingkat pembuangan limbah campuran selama pandemi COVID-19 dengan menggunakan variabel data sampah per hari, angka pengangguran, cuaca (suhu, kelembaban, dan kecepatan angin), dan kasus COVID-19 (pasien di rumah sakit, jumlah kasus, kasus aktif, dan uji baru) kemudian didapat lag dengan akurasi MAPE sebesar 9,4% sehingga LSTM dapat menghasilkan performa yang unggul meskipun dengan model sederhana dan variabel yang terbatas.[27]

No	Deskripsi
7.	Judul: <i>Detection of Long-Term Effect in Forecasting Municipal Solid Waste Using a Long Short-Term Memory Neural Network</i>
	Penulis: Dongjie Niu, Fan Wu, Shijin Dai, Sheng He, Boran Wu
	Tahun: 2020
	Data: Data harian timbulan sampah di Suzhou, China (2018-2019)
	Metode: LSTM, ARIMA, dan ANN konvensional
	Hasil: Penelitian ini dilakukan dengan metode LSTM untuk prediksi timbulan sampah padat kota (<i>municipal solid waste</i>) menggunakan variabel data historis MSW dan karakteristik waktu lalu dibandingkan juga dengan metode lain seperti ANN konvensional dan ARIMA didapat hasil akurasi LSTM tetap yang lebih unggul dari dataset <i>testing</i> didapat nilai RMSE 935,08, MAPE 114,36 dan R^2 0,92 sehingga LSTM lebih mampu untuk mendeteksi dalam jangka panjang dan menangkap pola dinamis dibanding metode pembanding.[28]
8.	Judul: <i>Bin Bay: An Optimized Smart Waste Disposal System for a Sustainable Urban Life Using LSTM and Fog</i>
	Penulis: Shiny Duella, Dioline Sara, Prabavathi
	Tahun: 2020
	Data: Sensor IoT dari smart bin di India (parameter: lokasi, kapasitas truk, status bin, dan rute)
	Metode: LSTM

No	Deskripsi
	Hasil: Penelitian yang dilakukan oleh Duela dkk pada tahun 2022 ini menggunakan metode <i>Long Short Term Memory</i> (LSTM) dan juga Fog Computing pada Bin Bay atau sistem pembuangan sampah cerdas yang diciptakan dengan menggunakan data jarak truk ke tempat pembuangan sampah, status untuk pengisian tempat sampah, lokasi truk, dan kapasitas truk untuk melakukan prediksi volume sampah per hari sehingga dapat meningkatkan efisiensi pengelolaan sampah.[29]
9.	Judul: <i>Comparative Analysis of SVM and LSTM in Municipal Solid Waste Management Models in Megacities</i>
	Penulis: Johanna Solano Meza, David Orjuela, Javier Rodrigo-Illari, Maria Elena Rodrigo-Clavero
	Tahun: 2023
	Data: Limbah padat yang ada di kota Bogota serta variabel eksternal.
	Metode: Support Vector Machines (SVM) dan LSTM
	Hasil : Penelitian ini dilakukan untuk dapat manajemen limbah padat yang ada di kota Bogota dengan menggunakan metode Long Short Term Memory (LSTM) dan membandingkannya dengan metode Support Vector Memory (SVM). Variabel yang digunakan meliputi produksi limbah, populasi, sosial-ekonomi, biaya transportasi, dan kemungkinan biaya pengolahan limbah. LSTM dinilai lebih unggul dalam pola data jangka panjang dibanding dengan SVM. Dan hasil penelitian menunjukkan skenario E6 paling efisien dari segi biaya dengan adanya potensi pendapatan tambahan dari produk sampingan.[30]

No	Deskripsi
10.	Judul: <i>Modeling of Municipal Waste Disposal Rates during COVID-19 Using Separated Waste Fraction Models</i>
	Penulis: Hoang Lan Vu, Kelvin Tsun Wai Ng, Amy Richter, Nima Karimi, Golam Kabir (University of Regina, Canada)
	Tahun: 2021
	Data: Data yang digunakan mencakup 7,5 tahun (dari Januari 2013 hingga September 2020), terdiri atas empat jenis sampah utama: sampah campuran (mixed), konstruksi dan pembongkaran (C&D), limbah medis (TBM), serta grit dan aspal.
	Metode: Long Short-Term Memory (LSTM)
	Hasil: Penelitian ini bertujuan untuk memodelkan jumlah pembuangan sampah kota (municipal solid waste/MSW) selama pandemi COVID-19 di kota Regina, Saskatchewan, Kanada. Peneliti membandingkan empat pendekatan pemodelan menggunakan model Long Short-Term Memory (LSTM), yaitu: (1) pendekatan total kontinyu (baseline, tanpa pemisahan jenis atau waktu), (2) pendekatan fraksi kontinyu (masing-masing jenis sampah dimodelkan secara terpisah lalu dijumlahkan), (3) pendekatan total terpotong (data total dipisah menjadi weekday dan weekend), dan (4) pendekatan fraksi terpotong (setiap fraksi sampah dipisahkan berdasarkan weekday dan weekend). Semua model dilatih menggunakan arsitektur LSTM dua lapis dengan 600 neuron dan lag waktu antara 10 hari hingga 2 minggu, tergantung jenis data. Hasil terbaiknya pada pemisahan berdasarkan fraksi dan waktu (weekday/weekend), yang berhasil menurunkan MAPE menjadi hanya 19,8% dan meningkatkan nilai korelasi R hingga 0,92.[31]

III. METODE PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian dan pengerjaan skripsi dilakukan pada:

Waktu : Desember 2024 – Mei 2025

Tempat : Universitas Lampung

Jadwal dan kegiatan penelitian yang dilakukan dapat dilihat pada Tabel 3.1

Tabel 3.1 Jadwal dan Kegiatan Penelitian

No.	Kegiatan	2024	2025				
		Des	Jan	Feb	Mar	Apr	Mei
1.	Studi Literatur dan Studi Pustaka						
2.	Persiapan Alat dan Bahan						
3.	Pemahaman Bisnis						
4.	Pemahaman Data						
5.	Persiapan Data						
6.	Pemodelan						
7.	Evaluasi						
8.	<i>Deployment</i>						

3.2 Alat dan Bahan Penelitian

3.2.1 Alat

Adapun alat yang digunakan dalam pengerjaan penelitian ini terdiri dari perangkat keras dan perangkat lunak.

- a. Perangkat keras yang digunakan Laptop dengan spesifikasi MSI dengan *processor* Intel Core i7-1195G7, RAM 8GB, SSD 512 GB, dan Windows 11 sebagai media merancang dan membuat algoritma prediksi berbasis LSTM.
- b. Perangkat lunak yang digunakan dapat dilihat pada Tabel 3.2

Tabel 3.2 Perangkat Lunak

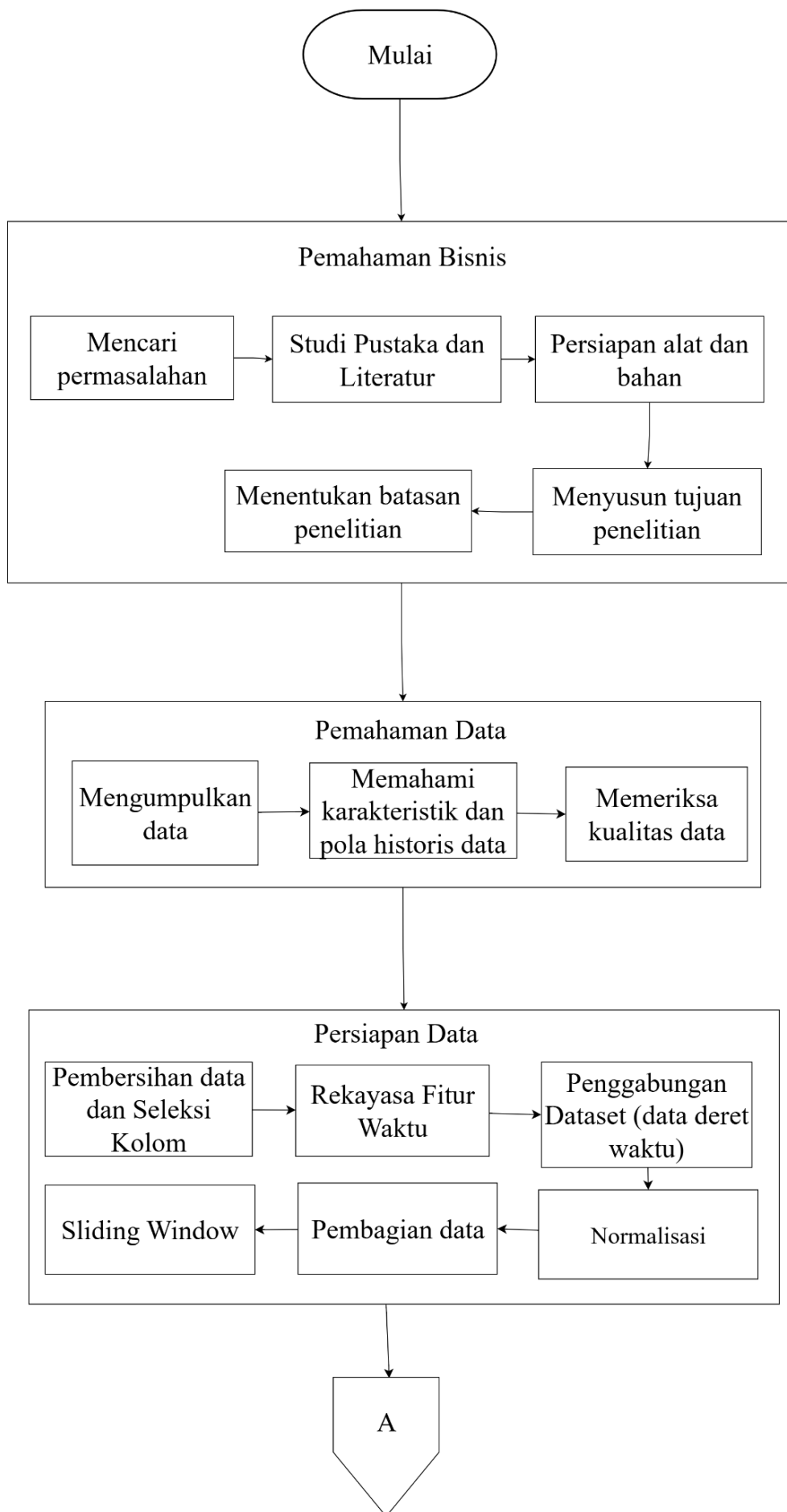
No.	Nama Alat	Deskripsi
1.	Microsoft Excel/ Google Sheets	Perangkat lunak yang digunakan dalam mengolah data mentah.
2.	Python 3.10.5	Bahasa pemrograman yang digunakan selama tahapan penelitian dengan metode CRISP-DM dilakukan.
3.	Google Collaboratory	Perangkat lunak berbasis web yang digunakan dalam menulis <i>syntax</i> dengan bahasa pemrograman python.
4.	Pandas	<i>Library</i> python yang digunakan untuk mengelola dan memproses dataset.
5.	<i>numpy</i>	<i>Library</i> python yang digunakan dalam operasi matematika dan manipulasi array seperti interpolasi atau normalisasi data.
6.	scikit-learn	<i>Library</i> yang digunakan untuk menghitung nilai metrik evaluasi seperti MAPE, RMSE, MAE, dan R^2 .
7.	TensorFlow	<i>Framework</i> yang digunakan dalam membangun dan melatih model LSTM.
8.	<i>matplotlib</i>	<i>Library</i> python yang digunakan untuk membuat visualisasi grafik sederhana.
9.	<i>seaborn</i>	<i>Library</i> python yang digunakan untuk membuat visualisasi statistik yang lebih kompleks.
10.	Streamlit	Perangkat lunak yang digunakan untuk menampilkan hasil prediksi dan tren dalam sebuah dashboard yang interaktif.

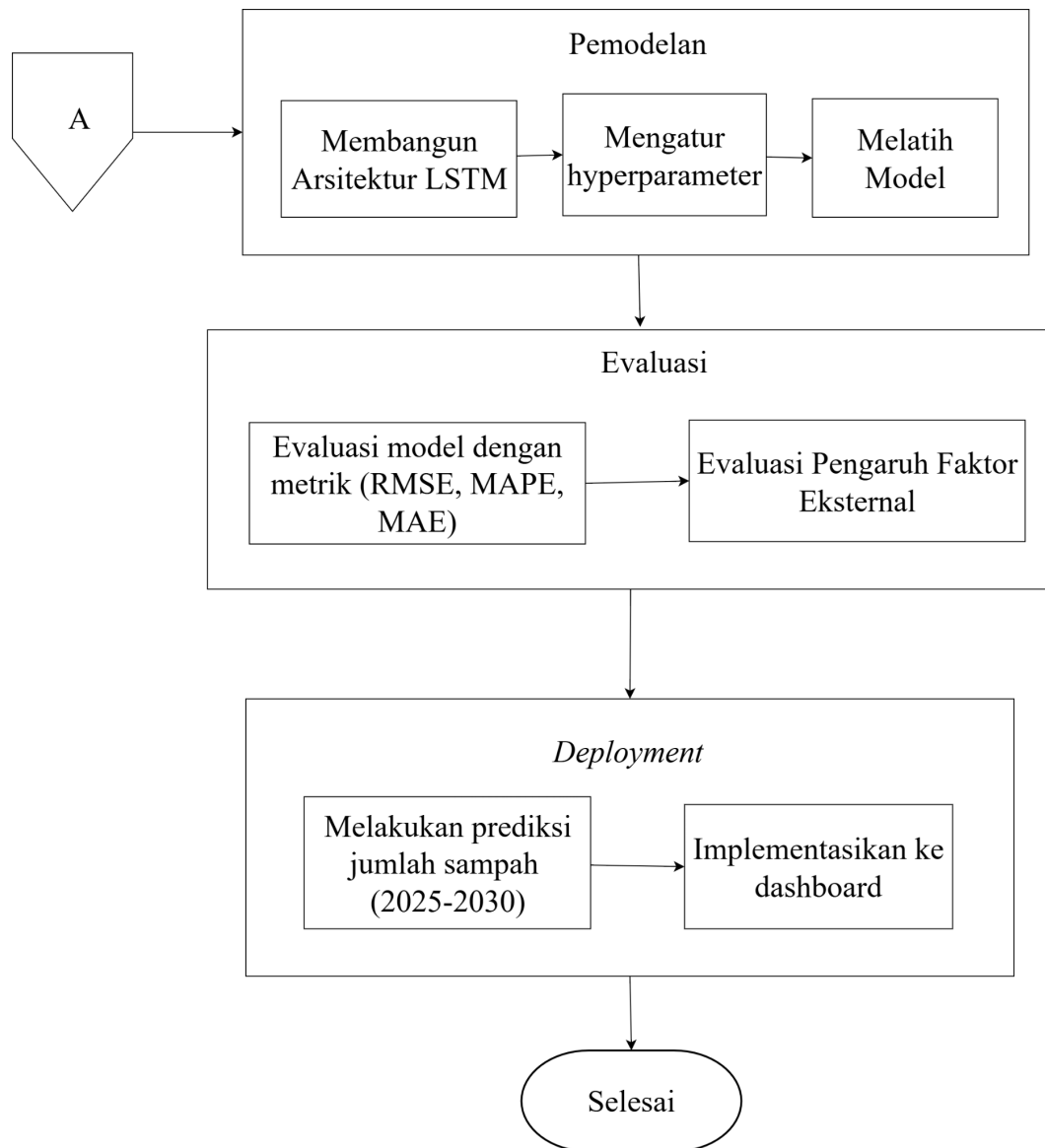
3.2.2 Bahan

Bahan yang digunakan pada penelitian ini adalah data sampah harian dengan rentang waktu per hari jumlah sampah per-kubik yang masuk pada TPA Bumi Ayu kabupaten Pringsewu tahun 2021 – 2024 yang diperoleh dari Dinas Lingkungan Hidup Kabupaten Pringsewu yang memiliki atribut pada data mentah seperti nomor, hari, tanggal, jumlah kendaraan pengangkut yang masuk (truk, pickup, dan roda tiga), volume yang masuk per kendaraan, lalu total seluruh kendaraan pengangkut sampah yang masuk, dan total seluruh volume sampah yang masuk. Selain itu terdapat juga data pendukung yaitu data jumlah penduduk per tahun Kabupaten Pringsewu yang didapat dari Data Kementerian Dalam Negeri dari tahun 2021 – 2024, data harian cuaca kabupaten Pringsewu dari tahun 2021 – 2024, dan data PDRB per kapita per tahun Kabupaten Pringsewu tahun 2021 – 2024 yang didapat dari Badan Pusat Statistik.

3.3 Tahapan Penelitian

Penelitian ini menggunakan metode CRISP-DM (*Cross - Industry Standard Process for Data Mining*) yang terdiri dari 6 tahap yaitu Pemahaman Bisnis, Pemahaman Data, Persiapan Data, Pemodelan, Evaluasi, dan *Deployment*. Tahapan ini disesuaikan dengan kebutuhan penelitian ini. Berikut diagram alir tahapan penelitian dapat dilihat pada Gambar 3.1





Gambar 3.1 *Flowchart* Tahapan Penelitian

3.3.1 Tahap Pemahaman Bisnis

Pada tahap pemahaman bisnis yang dilakukan pada penelitian ini bertujuan untuk mendapatkan pemahaman yang mendalam tentang konteks, masalah, tujuan, serta batasan dari penelitian yang akan dilakukan sehingga dapat disusun alur dan strategi yang jelas sebelum memasuki tahap teknis. Tahap ini dimulai dari pencarian ide dan topik penelitian yang dilakukan melalui diskusi dengan dosen pembimbing dan pengamatan terhadap permasalahan aktual atau isu di

lingkungan. Dari hasil pengamatan lapangan dan menelaah berita lokal peneliti memilih fokus pada isu pengelolaan sampah di Kabupaten Pringsewu setelah didapatkan jika pada TPA Bumiayu berpotensi menghadapi masalah kapasitas dalam beberapa tahun ke depan, meskipun pihak terkait sudah memiliki rencana terkait pembukaan lahan TPA baru. Berdasarkan situasi tersebut ditemukan peluang untuk melakukan penelitian berbasis kecerdasan buatan untuk memprediksi jumlah sampah di masa depan sebagai bahan perencanaan pengelolaan sampah.

Untuk menguatkan dasar penelitian ini, dilakukan studi pustaka dan literatur yang mencakup penelusuran dengan kata kunci yang relevan untuk membantu pencarian literatur terkait dan didapat 10 jurnal internasional terkait, lalu didapat juga 16 referensi buku dan 6 sumber dari yang membahas metode peramalan data deret waktu, pengelolaan sampah, serta model prediksi. Studi pustaka ini juga dilakukan untuk mendapatkan pemahaman mendalam terkait penelitian terdahulu, kelebihan dan kekurangan metode, serta langkah terbaik yang dapat digunakan untuk penelitian ini. Maka setelah dilakukan pencarian melalui internet dilakukan diskusi kembali dengan dosen pembimbing untuk metode serta variabel yang digunakan. Berdasarkan referensi yang digunakan maka diperoleh metode LSTM dengan variabel yang berpotensi mempengaruhi jumlah sampah seperti data cuaca, jumlah penduduk, dan indikator ekonomi yang dipertimbangkan berdasarkan penelitian-penelitian terkait yang ditemukan.

Kemudian dilakukan persiapan alat dan bahan penelitian. Data utama yang digunakan adalah data volume sampah harian periode 2021 – 2024 yang diperoleh dari Dinas Lingkungan Hidup (DLH) Kabupaten Pringsewu melalui proses diskusi penyampaian rencana penelitian dan perizinan resmi dengan surat permohonan data kepada Kepala Bidang Pengelolaan Sampah, Limbah B3, dan Pencemaran untuk memperoleh informasi terkait pengelolaan data, kapasitas harian, dan proyeksi pengelolaan sampah di masa depan yang memperkuat urgensi penelitian untuk mengembangkan model prediksi jumlah sampah berbasis LSTM sehingga dapat mencapai tujuan penelitian yang meliputi:

- a. Mengembangkan sistem kecerdasan buatan dengan model LSTM yang mampu memprediksi jumlah sampah berdasarkan data historis jumlah sampah harian di TPA Bumiayu.
- b. Menganalisis performa model LSTM dalam memprediksi jumlah sampah harian berdasarkan data historis sampah di TPA Bumiayu.
- c. Menganalisis model LSTM jika dilakukan penambahan fitur variabel pengaruh faktor cuaca, jumlah penduduk, dan pendapatan di Kabupaten Pringsewu dalam memprediksi jumlah sampah di TPA Bumiayu.
- d. Mengembangkan dashboard visualisasi hasil prediksi jumlah sampah setelah diimplementasikan dengan model LSTM selama 2025-2030 dari TPA Bumiayu.

Sedangkan batasan penelitian untuk permasalahan yang diambil yaitu

- a. Targetnya adalah memprediksi jumlah sampah dalam dari tahun 2025-2030.
- b. Data yang digunakan adalah data historis sampah di TPA Bumiayu yang dikelola oleh Dinas Lingkungan Hidup Kabupaten Pringsewu yang terkumpul sejak tahun 2021-2024.
- c. Untuk fitur eksternal yang digunakan untuk melihat pengaruhnya terhadap jumlah sampah yang masuk di Kabupaten Pringsewu meliputi data cuaca yang diambil dari Visual Crossing yang hanya akan dipilih 1-2 fitur lalu ada data jumlah penduduk dan PDRB Per Kapita yang diambil dari Badan Pusat Statistik Kabupaten Pringsewu dimana akan diuji setelah mendapat model terbaik.
- d. Model yang digunakan adalah berbasis LSTM sebagai model untuk mengetahui akurasi model, stabilitas model, serta tren aktual dan prediksi terhadap targetnya yaitu jumlah sampah dengan menggunakan Single layer, Double layer, dan Bidirectional layer serta melakukan kombinasi hyperparameter menggunakan epoch antara 50 dan 100, *neuron* antara 64 dan 128, serta learning rate antara 0,001 dan 0,0005. Serta menggunakan sliding window yaitu 30.

- e. Melakukan evaluasi model dengan menggunakan indikator metrik yaitu *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), dan *Root Mean Square Error* (RMSE).
- f. Melakukan prediksi jangka panjang dengan model terbaik baik berdasarkan akurasi, stabilitas proses pelatihan (grafik training dan validation loss), dan kualitas visualisasi prediksi terhadap data aktual.
- g. Melakukan visualisasi hasil prediksi dengan menggunakan Streamlit untuk *long term prediction* dengan model terbaik.

3.3.2 Tahap Pemahaman Data

Pada tahap pemahaman data ini setelah mendapatkan izin dari pihak dinas lingkungan hidup Kabupaten Pringsewu atau staf pengelola TPA dilakukan pengamatan terhadap pola, struktur dan kualitas data agar lebih mudah dipahami. Untuk data sampah yang diambil ada data volume sampah harian dari TPA Bumiayu sejak tahun 2021 – 2024 yang akan dijadikan sebagai target, lalu data jumlah penduduk dan PDRB per kapita Kabupaten Pringsewu yang diambil dari Badan Pusat Statistik Pringsewu, dan data cuaca harian untuk wilayah Pringsewu sejak tahun 2021-2024 dari situs Visual Crossing. Pada data mentah sampah harian yang didapat ini mencakup atribut seperti nomor, hari, tanggal, jumlah kendaraan pengangkut yang masuk (truk, pickup, dan roda tiga), volume yang masuk per kendaraan, lalu total seluruh kendaraan pengangkut sampah yang masuk, dan total seluruh volume sampah yang masuk dengan informasi dan istilah untuk atribut dalam data didapat dari staf pengelola TPA.

Kemudian untuk data cuaca yang dapat dikumpulkan mencakup atribut seperti suhu minimum, suhu maksimum, suhu rata-rata, suhu terasa minimum, suhu terasa maksimum, suhu terasa rata-rata, titik embun, curah hujan, persentase wilayah hujan, hembusan angin, kecepatan angin, arah angin, tutupan awan, kelembaban harian, radiasi matahari, dan energi matahari dari data mentah yang didapat. Sementara ada juga data pendapatan Produk Domestik Regional Bruto (PDRB) Per Kapita Atas Dasar Harga Konstan (ADHK) yaitu harga nilai barang

dan jasa dari suatu wilayah dalam periode tertentu yang diambil berdasarkan harga yang berlaku pada tahun tertentu sehingga dapat membandingkan pertumbuhan ekonomi antar tahun tanpa terpengaruh inflasi. Dan data jumlah penduduk per tahun sejak 2021 – 2024. Untuk data sampah dan cuaca memiliki 1.461 baris data. Berikut sampel data mentah yang telah dikumpulkan cuplikan ini berupa gambaran yang dapat dilihat untuk mengetahui struktur dan atribut data bukan secara keseluruhan isi dataset.

Tabel 3.3 Sampel Data Jumlah Sampah TPA Bumiayu Tahun 2021-2024

NO.		1	2	3	4	5
HARI		Jumat	Sabtu	Minggu	Senin	Selasa
TANGGAL		01/01/ 2021	02/01/ 2021	03/01/ 2021	04/01/ 2021	05/01/ 2021
KENDARAAN	TRUK	8	8		8	9
	VOL 8M ³	64	64		64	72
	PICK UP	1	2		1	2
	VOL 4 M ³	4	8		4	8
	RODA 3	6	5		6	5
	VOL 1,5 M ³	9	7.5		9	7.5
TOTAL	KENDARAAN	15	15		15	16
	VOL. SELURUH M ³	77	79,5		77	87,5

Sumber: Bagian Administrasi TPA Bumiayu

Tabel 3.4 Data Jumlah Penduduk dan PDRB Per Kapita Kabupaten Pringsewu

Tahun	Jumlah Penduduk	PDRB Per Kapita (Rp)
		ADHK
2021	409330	19.226
2022	414470	19.818
2023	419590	20.513
2024	424680	21.139

Sumber: Badan Pusat Statistik Pringsewu

Data jumlah penduduk dan PDRB per kapita akan di resample menjadi harian agar memiliki format data yang sama.

Tabel 3.5 Sample Data Cuaca Kabupaten Pringsewu

Lokasi	Pringsewu	Pringsewu	Pringsewu	Pringsewu	Pringsewu
Tanggal	01/01/2021	02/01/2021	03/01/2021	04/01/2021	05/01/2021
Suhu Maksimum (C)	31,7	31,1	31,6	30,4	30,4
Suhu Minimum (C)	22,9	22,9	23,5	23,2	24,2
Suhu Rata-Rata (C)	26,2	25,9	25,6	26	26,4
Suhu Terasa Maksimum (C)	37,3	37	37,1	34,8	35,8
Suhu Terasa Minimum (C)	22,9	22,9	23,5	23,2	24,2
Suhu Terasa Rata-Rata (C)	27,5	27	26,8	27,2	27,8
Titik Embun (C)	23,3	23,3	23,3	23,5	23,9
Kelembaban Relatif (%)	84,9	86,5	87,7	86,8	86,5
Curah Hujan (mm)	0,4	5,8	42	4	0,3
Persentase Wilayah Hujan (%)	8,33	16,67	4,17	41,67	8,33
Hembusan Angin (km/jam)	29,2	27,4	22,7	25,9	38,5
Kecepatan Angin (km/jam)	9,4	9,4	7,6	5,4	13,3
Arah Angin (derajat)	321,9	314,8	230,5	293,8	340,5
Tutupan Awan (%)	90,3	98,5	98,1	96,6	96,6
Jarak Pandang	8,8	7,6	9,6	9,1	8,4
Radiasi Matahari (m2)	233,9	171,8	186,7	141,6	198,1
Energi Matahari (W/m2)	20,2	14,8	16	12,2	17

Sumber: Visual Crossing

Pada tahap eksplorasi data dari data yang telah dikumpulkan dilakukan analisis deskriptif untuk melihat relevansi atribut yang ada untuk prediksi jumlah sampah ini dilakukan dengan membuat visualisasi untuk melihat hubungan antar variabel.

Selain itu juga pada tahap ini data akan diperiksa kualitasnya dengan melihat dari visualisasi yang dilakukan ditemukan atau tidak data duplikat atau nilai yang kosong (*missing values*) atau terdapat format yang tidak sesuai pada data.

3.3.3 Tahap Persiapan Data

Pada tahap persiapan data dilakukan beberapa tahapan untuk mempersiapkan data yang siap digunakan pada tahap pemodelan yaitu

- 1) Data yang sudah dikumpulkan berupa sampah harian, jumlah penduduk, cuaca, dan PDRB per kapita. Untuk data jumlah penduduk dan PDRB per kapita dilakukan resampling data dengan *forward resampling*, yaitu nilai tahun tertentu digunakan tetap sepanjang tahun bersangkutan sehingga formatnya menjadi harian mengikuti frekuensi data target setelah itu dilakukan analisis deskriptif dan visualisasi untuk melihat kualitas data maka data akan dibersihkan dari *missing values* dan *outlier* yang terdeteksi pada saat tahap pemahaman data dilakukan.
- 2) Dibangun fitur rekayasa waktu secara eksplisit dalam pemodelan untuk membantu LSTM mempelajari urutan data secara mendalam agar dapat menangkap pola musiman atau perulangan yang ada.
- 3) Setelah didapat atribut yang sesuai maka data-data tersebut akan digabungkan yang disusun berdasarkan atribut tanggal agar membentuk dataset yang konsisten sehingga siap untuk model prediksi sesuai dengan kebutuhan model LSTM yaitu data dengan urutan waktu berkala (*time series*).
- 4) Data akan dilakukan *sliding window* yaitu metode yang memotong data deret waktu menjadi beberapa *window* (jendela) berukuran tetap yaitu 30, dengan tujuan agar model dapat mempelajari pola berdasarkan urutan data historis. Metode ini dilakukan secara bertahap untuk menghasilkan *input (features)* dan *output (labels)* yang mana dibuat untuk memudahkan dalam membaca pola baik harian, bulanan, ataupun tahun.
- 5) Dilakukan pembagian data untuk membagi data menjadi 70% untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian.

- 6) Setelah itu dinormalisasi ke rentang nilai antara 0 dan 1 dengan metode Min-Max Scaling supaya model dapat mengolah input lebih stabil.

3.3.4 Tahap Pemodelan

Pada tahap pemodelan, dilakukan eksperimen menggunakan arsitektur Long Short-Term Memory (LSTM) untuk membangun model baseline dalam memprediksi jumlah timbulan sampah harian. Eksperimen dilakukan terhadap beberapa struktur jaringan, seperti single-layer LSTM, double-layer LSTM, dan bidirectional LSTM (Bi-LSTM), dengan kombinasi *hyperparameter* seperti jumlah neuron, jumlah epoch, ukuran *batch*, dan *learning rate*. Proses ini dilakukan untuk memperoleh konfigurasi model yang optimal dan stabil. Setiap kombinasi dilatih menggunakan data pelatihan dan divalidasi pada data validasi, dengan pemantauan terhadap grafik *loss* untuk mendeteksi gejala *overfitting* atau *underfitting*.

3.3.5 Tahap Evaluasi

Setelah pelatihan, setiap model dievaluasi menggunakan metrik Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), dan Mean Absolute Percentage Error (MAPE), guna mengukur tingkat akurasi dan kesalahan prediksi. Pemilihan model terbaik ditentukan berdasarkan performa terbaik pada data latih lalu kestabilan pada data validasi dan juga akurasi prediksi pada data uji. Selanjutnya, arsitektur dan konfigurasi model terbaik ini digunakan untuk mengevaluasi pengaruh penambahan fitur eksternal seperti cuaca, ekonomi, dan kependudukan dalam pendekatan multivariate, guna mengamati peningkatan akurasi prediksi secara adil dan terukur.

3.3.6 Tahap Penyebaran (*Deployment*)

Tahap terakhir adalah tahap *deployment* dengan melakukan prediksi sampah harian tahun 2025 – 2030 dengan fitur dan model terbaik, hasil prediksi akan digunakan untuk membuat dashboard visualisasi dengan menampilkan hasil prediksi dalam bentuk grafik yang dibuat dengan *library* matplotlib dan plotly.

Dashboard akan dibuat menggunakan Streamlit dipilih karena kebutuhan yang sederhana, gratis, mudah digunakan sehingga juga mudah diintegrasikan meskipun pengolahan data dan pemodelan dilakukan pada Google Collaboratory.

V. PENUTUP

5.1 Kesimpulan

Adapun beberapa kesimpulan penelitian ini dari efektivitas dan akurasi model LSTM dalam memprediksi data jumlah sampah dari TPA Bumiayu lalu mengetahui pengaruh faktor eksternal dan hasil prediksi jumlah sampah di TPA 2025 – 2030 dengan model terbaik.

1. Hasil terbaik yang dikembangkan dengan model LSTM untuk memprediksi jumlah sampah harian di TPA Bumiayu dengan data historis 2021 – 2024 dan diperoleh dari arsitektur single layer menggunakan 128 neuron, 100 epoch, learning rate 0,0005, dan batch size 64 menghasilkan MAPE sebesar 6,37%, model yang stabil, dan grafik mengikuti tren aktual.
2. Dari hasil pelatihan model berbasis LSTM dengan menggunakan arsitektur Single, Double, dan Bidirectional dengan kombinasi hyperparameter menggunakan kombinasi target dan fitur waktu menunjukkan arsitektur single layer memiliki akurasi terbaik sedangkan arsitektur double dan bidirectional tidak memberikan peningkatan performa.
3. Penambahan fitur dari variabel eksternal seperti curah hujan, jumlah penduduk, dan PDRB per kapita setelah diperoleh model LSTM terbaik tidak meningkatkan performa model secara signifikan hanya curah hujan yang performanya stabil meskipun MAPE menjadi 6,41% hal ini menunjukkan bahwa tidak semua variabel eksternal relevan.
4. Prediksi jumlah sampah harian tahun 2025 – 2030 menunjukkan tren kenaikan bertahap dengan pola musiman yang konsisten setiap tahun. Dari hasil prediksi yang divisualisasikan dalam bentuk dashboard meskipun dengan keterbatasan mendeteksi perubahan skala data dapat mendukung perencanaan jangka panjang atau pembentukan regulasi di TPA Bumiayu

dengan didukung perubahan secara berkala untuk perbaikan pengelolaan sampah.

5.2 Saran

Adapun saran penelitian untuk mendukung pengembangan penelitian ini dilakukan.

1. Disarankan untuk mengeksplorasi fitur eksternal tambahan yang memiliki hubungan kausal langsung terhadap volume sampah, seperti data aktivitas masyarakat, kalender event lokal, atau data operasional pengangkutan. Fitur ini dapat membantu model mengenali potensi lonjakan volume sampah non musiman dan jangka panjang, namun, penambahan fitur harus dilakukan secara selektif agar tidak mengganggu pola mayoritas dan memperburuk performa model LSTM yang sensitif terhadap fluktuasi data.
2. Salah satu keterbatasan LSTM adalah dalam menangkap lonjakan ekstrem yang tidak teratur, disarankan pendekatan hybrid seperti menggabungkan model statistik (misalnya Prophet) dengan model neural (seperti LSTM) mungkin dapat menjadi solusi. Prophet untuk memodelkan tren dan pola musiman, sementara LSTM dilatih pada residual-nya untuk menangkap komponen non-musiman dan anomali.
3. Salah satu langkah yang dapat dipertimbangkan untuk meningkatkan akurasi prediksi adalah penerapan teknologi *tagging* pada armada pengangkut sampah seperti RFID dan GPS untuk membantu merekam identitas kendaraan, waktu masuk TPA, dan volume sampah secara otomatis. Dengan integrasi ke dalam sistem pencatatan berbasis database membantu menyediakan informasi harian secara real-time dan mengurangi risiko kesalahan input manual serta mendukung pengembangan prediksi berbasis *real-time forecasting*.
4. Untuk menjaga akurasi dan relevansi prediksi jangka panjang (seperti untuk tahun 2025–2030), model sebaiknya secara berkala diperbarui menggunakan data terbaru yang terus bertambah seiring waktu. Pembaruan ini penting

untuk memastikan bahwa model tetap adaptif terhadap perubahan pola dan tren yang mungkin terjadi akibat dinamika sosial, ekonomi, atau lingkungan.

5. Hasil prediksi dapat dimanfaatkan dalam penyusunan rencana pengelolaan sampah daerah, terutama untuk merencanakan kapasitas TPA, logistik pengangkutan, armada, SDM, dan anggaran jangka panjang. Meskipun, akurasi prediksi masih dipengaruhi oleh keterbatasan data historis yang tersedia, penguatan sistem pencatatan data serta sinergi antara pemerintah, akademisi, dan masyarakat menjadi langkah penting dalam membangun pengelolaan sampah yang lebih adaptif dan berbasis data. Selain itu, peningkatan metode pengelolaan teknis sampah yang lebih baik juga penting untuk meminimalkan potensi permasalahan serupa di masa mendatang meskipun infrastruktur baru telah disiapkan.

DAFTAR PUSTAKA

- [1] “Laporan Kumpulan Data Demografi Penduduk Indonesia Semester I 2024.” Accessed: Dec. 22, 2024. [Online]. Available: <https://assets.dataindonesia.id/2024/09/02/1725256436251-24-Laporan-Kumpulan-Data-Demografi-Penduduk-Indonesia-Semester-I-2024.pdf>
- [2] “SIPSN - Sistem Informasi Pengelolaan Sampah Nasional.” Accessed: Dec. 22, 2024. [Online]. Available: <https://sipsn.menlhk.go.id/sipsn/>
- [3] A. A. Althusius, “Paparan TPA Bumiayu,” Scribd. Accessed: July 07, 2025. [Online]. Available: <https://id.scribd.com/presentation/668335446/Paparan-TPA-Bumiayu>
- [4] F. Chollet, *Deep Learning with Python (2nd Edition)*. Shelter Island: Manning Publications, 2021.
- [5] J. Brownlee, “Long Short-Term Memory Networks With Python”.
- [6] M. Cubillos, “Multi-Site Household Waste Generation Forecasting Using a Deep Learning Approach,” *Waste Manag.*, vol. 115, pp. 8–14, Sept. 2020, doi: 10.1016/j.wasman.2020.06.046.
- [7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [8] Dr. C. Chotimah, *Pengelolaan Sampah dan Pengembangan Ekonomi Kreatif di Kawasan Destinasi Wisata Pesisir Pantai Selatan Tulungagung: Community Engagement pada Masyarakat Klatak dan Keboireng Besuki Tulungagung*. Tulungagung: Akademia Pustaka, 2020.
- [9] A. Suprihatin, D. Prihanto, and M. Gelbert, *Buku Panduan*. Malang: PPPGT/VEDC Malang, 1999.

- [10] N. L. W. S. R. Ginantra *et al.*, *Data Mining Dan Penerapan Algoritma*. Denpasar: Yayasan Kita Menulis, 2021.
- [11] H. Jaya, Sabran, and M. Idris, *Kecerdasan Buatan*. Makassar: Fakultas MIPA Universitas Negeri Makassar, 2018.
- [12] N. Yudistira, *Deep Learning: Teori, Contoh, Perhitungan, dan Implementasi*. Yogyakarta: Deepublish, 2024.
- [13] I. Cholissodin, *AI, Machine Learning & Deep Learning (Teori & Implementasi)*. Malang: Fakultas Ilmu Komputer Universitas Brawijaya, 2020.
- [14] A. Khumaidi and I. A. Nirmala, *Algoritma Long Short Term Memory Dengan Hyperparameter Tuning: Prediksi Penjualan Produk*. Yogyakarta: Deepublish, 2022.
- [15] N. Yudistira, L. M. D. Alfiansih, and N. I. Andriyani, *Prediksi Deret Waktu Menggunakan Deep Learning*. Malang: UB Press, 2023.
- [16] Drs. J. J. Siang, *Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan Matlab*. Yogyakarta: ANDI, 2005.
- [17] A. P. Windarto *et al.*, *Jaringan Saraf Tiruan: Algoritma Prediksi dan Implementasi*. Medan: Yayasan Kita Menulis, 2020.
- [18] H. A. Khoiri, *Analisis Deret Waktu Univariat (Teori dan Pengolahan Data)*. Madiun: UNIPMA Press, 2023.
- [19] Z. Sitorus, *Panduan Praktis Analisis Statistik Untuk Penelitian Skripsi, Thesis, dan Disertasi*, Cetakan I. Medan: PT Media Penerbit Indonesia, 2024.
- [20] J. Walkenbach, *Excel 2016 Bible*. Indianapolis: John Wiley & Sons, Inc., 2015.
- [21] “Google Colab.” Accessed: Dec. 22, 2024. [Online]. Available: <https://colab.research.google.com/>
- [22] “Streamlit Docs.” Accessed: June 09, 2025. [Online]. Available: <https://docs.streamlit.io/>

- [23] A. Suharto, *Fundamental Bahasa Pemrograman Python*. Purbalingga: Eureka Media Aksara, 2021.
- [24] M. A. Olawumi, B. I. Oladapo, and R. A. Olawale, “Revolutionising Waste Management with the Impact of Long Short-Term Memory Networks on Recycling Rate Predictions,” *Waste Manag. Bull.*, vol. 2, no. 3, pp. 266–274, Sept. 2024, doi: 10.1016/j.wmb.2024.08.006.
- [25] Department of Computer Engineering, Munzur University, Tunceli, Turkey, A. Utku, S. K. Kaya, and Department of Industrial Engineering, Munzur University, Tunceli, Turkey, “Deep Learning Based a Comprehensive Analysis for Waste Prediction,” *Oper. Res. Eng. Sci. Theory Appl.*, vol. 5, no. 2, pp. 176–189, Aug. 2022, doi: 10.31181/oresta190822135u.
- [26] L. Huang, T. Cai, Y. Zhu, Y. Zhu, W. Wang, and K. Sun, “LSTM-Based Forecasting for Urban Construction Waste Generation,” *Sustainability*, vol. 12, no. 20, p. 8555, Oct. 2020, doi: 10.3390/su12208555.
- [27] H. L. Vu, K. T. W. Ng, A. Richter, and G. Kabir, “The Use of a Recurrent Neural Network Model with Separated Time-Series and Lagged Daily Inputs for Waste Disposal Rates Modeling During COVID-19,” *Sustain. Cities Soc.*, vol. 75, p. 103339, Dec. 2021, doi: 10.1016/j.scs.2021.103339.
- [28] D. Niu, F. Wu, S. Dai, S. He, and B. Wu, “Detection of Long-Term Effect in Forecasting Municipal Solid Waste Using a Long Short-Term Memory Neural Network,” *J. Clean. Prod.*, vol. 290, p. 125187, Mar. 2021, doi: 10.1016/j.jclepro.2020.125187.
- [29] S. Duella, D. Sara, and Prabavathi, “Bin Bay: An Optimized Smart Waste Disposal System for a Sustainable Urban Life using LSTM and Fog Computing,” *J. Comput. Sci.*, vol. 18, no. 11, pp. 1110–1120, Nov. 2022, doi: 10.3844/jcssp.2022.1110.1120.
- [30] J. K. Solano Meza, D. Orjuela Yepes, J. Rodrigo-Ilarri, and M.-E. Rodrigo-Clavero, “Comparative Analysis of the Implementation of Support Vector Machines and Long Short-Term Memory Artificial Neural Networks in Municipal Solid Waste Management Models in Megacities,” *Int. J.*

Environ. Res. Public. Health, vol. 20, no. 5, p. 4256, Feb. 2023, doi: 10.3390/ijerph20054256.

- [31] H. L. Vu, K. T. W. Ng, A. Richter, N. Karimi, and G. Kabir, “Modeling of Municipal Waste Disposal Rates during COVID-19 Using Separated Waste Fraction Models,” *Sci. Total Environ.*, vol. 789, p. 148024, Oct. 2021, doi: 10.1016/j.scitotenv.2021.148024.
- [32] S. Kapoor and A. Narayanan, “Leakage and the Reproducibility Crisis in Machine-Learning-Based Science,” *Patterns*, vol. 4, no. 9, p. 100804, Sept. 2023, doi: 10.1016/j.patter.2023.100804.