# DAFTAR ISI

[DAFTAR ISI i](#_Toc141243272)

[BAB I PENDAHULUAN 1](#_Toc141243273)

[1.1 Latar Belakang 1](#_Toc141243274)

[1.2 Rumusan Masalah 2](#_Toc141243275)

[1.3 Batasan Masalah 2](#_Toc141243276)

[1.4 Tujuan Penelitian 3](#_Toc141243281)

[1.5 Manfaat Penelitian 3](#_Toc141243282)

[BAB II TINAJAUAN PUSTAKA 5](#_Toc141243283)

[2.1 Research Puzzle 5](#_Toc141243284)

[2.2 Tinjauan Pustaka 5](#_Toc141243285)

[2.3 Landasan Teori 6](#_Toc141243286)

[BAB III METODE PENELITIAN 11](#_Toc141243287)

[3.1 Bentuk Penelitian 11](#_Toc141243290)

[3.2 Pengumpulan Data 11](#_Toc141243291)

[3.3 Tahapan Penelitian 11](#_Toc141243292)

[3.3.1 Business Understanding 12](#_Toc141243299)

[3.3.2 Data Understanding 12](#_Toc141243300)

[3.3.3 Data Preparation 12](#_Toc141243301)

[3.3.4 Modeling 12](#_Toc141243302)

[3.3.5 Evaluation 12](#_Toc141243303)

[3.3.6 Deployment 13](#_Toc141243304)

[BAB IV HASIL DAN PEMBAHASAN 15](#_Toc141243305)

[4.1 Data 15](#_Toc141243309)

[4.2 Preprocessing Data 15](#_Toc141243310)

[4.3 Pemodelan dengan LSTM 16](#_Toc141243311)

[4.4 Pemodelan dengan RNN 17](#_Toc141243312)

[4.5 Pemodelan dengan RNN-LSTM 18](#_Toc141243313)

[4.6 Pemodelan dengan LSTM-RNN 18](#_Toc141243314)

[4.7 Matriks Evaluasi (MAE & *LOSS*) 19](#_Toc141243315)

[4.8 Hasil Prediksi (MAE) 25](#_Toc141243316)

[4.9 Conifdence Interval 27](#_Toc141243317)

[BAB V KESIMPULAN DAN SARAN 40](#_Toc141243318)

[5.1 Kesimpulan 40](#_Toc141243322)

[5.2 Saran 40](#_Toc141243323)

[DAFTAR PUSTAKA 42](#_Toc141243324)

# BAB I PENDAHULUAN

## Latar Belakang

Asset Integrity Management (AIM) telah mengalami perkembangan selama beberapa dekade terakhir, mulai dari inspeksi berbasis waktu yang sederhana pada peralatan kunci hingga sistem manajemen berbasis risiko dan keandalan untuk semua bagian yang kritis terhadap keselamatan (AMAEchi, 2022). Kerusakan atau kegagalan aset dapat mengakibatkan kecelakaan, gangguan operasional, penurunan produksi, dan bahkan dampak lingkungan yang serius (Earthjustice, 2022). Dalam industri minyak dan gas, aset seringkali beroperasi di kondisi yang ekstrem, seperti tekanan tinggi, suhu ekstrim, dan lingkungan korosif. Hal ini meningkatkan risiko korosi, kelelahan material, retak, dan penurunan kinerja aset seiring waktu.

Tradisionalnya, pemeliharaan aset dilakukan melalui inspeksi manual dan jadwal pemeliharaan rutin. Namun, pendekatan ini dapat menjadi mahal, memakan waktu, dan sulit untuk mendeteksi masalah yang tersembunyi atau berkembang secara bertahap. Oleh karena itu, telah banyak peneltian yang mulai menerapkan kemampuan Artificial Intelligent dalam melakukan manejemen pada sebuah Asset Integrity, seperti penenlitian dibidang oil dan gas oleh Sun et al. (2019), Yang et al. (2020), Zhou et al. (2021), dan Maythalony et al. (2019). Kemudian, penelitian sejenis akan diusulkan dengan sebutan Intelligent Asset Integrity (IAI).

Tujuan utama dari Intelligent Asset Integrity adalah meningkatkan deteksi dan identifikasi kesalahan secara real-time melalui ramalan kondisi sistem di masa depan dalam rentang waktu tertentu. Keuntungan utama dari proses ini adalah pemicu layanan cepat yang mencegah waktu henti (Ossae, 2017). Dengan menggunakan IAI, perusahaan dapat mengumpulkan dan menganalisis data aset secara *real-time*, memperkirakan keandalan dan masa pakai sisa aset, serta mengidentifikasi indikasi awal potensi kegagalan atau masalah integritas. Hal ini memungkinkan pengambilan keputusan yang lebih cepat dan tepat dalam melakukan tindakan perbaikan atau pemeliharaan, mengurangi risiko kegagalan, mengoptimalkan kerja pemantauan, dan memaksimalkan efisiensi operasional.

Kegagalan dalam menjaga Asset Integrity dapat menyebabkan kerugian yang sangat besar bahkan korban jiwa, terutama dalam industri minyak dan gas. Di Indonesia, berdasarkan data dari CNBC, tercatat bahwa pada tanggal 3 Maret 2023, terjadi insiden di Depo Plumpang yang berlokasi di Jakarta Utara. Hal ini mengakibatkan 17 orang meninggal dan puluhan orang luka (Pgr, 2023). Pada tahun yang sama, tanggal 1 April 2023, terjadi ledakan di Kilang Bahan Bakar Minyak yang berlokasi di kota Dumai provinsi Riau, ledakan ini menewaskan 5 orang korban jiwa (Asmarini, 2023). Kepolisian Riau melaporkan bahwa penyebab ledakan tersebut adalah kebocoran gas hidrogen (H2) yang kemudian diikuti oleh terbakarnya Hydro Cracker (HCU) (Diskominfotik Riau, 2023).

Berdasarkan latar belakang tersebut, pentingnya Asset Integrity dalam industri minyak dan gas menjadi semakin signifikan. Melalui penerapan kecerdasan buatan atau Intelligent Asset Integrity (IAI), perusahaan dapat menghadirkan solusi yang efektif dalam meningkatkan efisiensi kerja serta mengurangi biaya operasional. Dengan memanfaatkan IAI, perusahaan dapat meningkatkan tingkat keamanan, keandalan, efisiensi, dan keberlanjutan operasional aset mereka. Dengan demikian, penggunaan teknologi kecerdasan buatan dalam Asset Integrity memberikan manfaat yang signifikan bagi perusahaan, mendukung pengambilan keputusan yang lebih baik, serta membantu mereka mencapai tujuan operasional mereka dengan lebih efektif dan efisien.

Penelitian terhadap Intelligent Asset Integrity telah banyak dilakukan. Beberapa diantaranya yaitu, Sun et al (2019) menggunakan model Artificial Neural Network (ANN) untuk memprediksi sisa usia pakai dari bantalan. Yang et al. (2020) menggunakan Convolutional Neural Networks (CNN) untuk mendeteksi kerusakan pada bilah turbin angin berdasarkan *computer vision*. Zhou et al. (2021) menggunakan data rentang waktu dan metode Long Short-Term Memory (LSTM) untuk memprediksi sisa usia pakai dari bilah turbin angin. Al-Maythalony et al. (2019), menggunakan Random Forest (RF) untuk memprediksi laju korosi yang terjadi pada aset gas minyak dan gas. Selain itu, terdapat penelitian Setiawan et al (2021) yang menggunakan analisis regresi untuk menemukan variabel dengan efek signifikan pada usia pakai lapisan.

Penelitian yang dilakukan oleh Sun et al. (2019), Yang et al. (2020), Zhou et al. (2021), dan Maythalony et al. (2019) telah menghasilkan model-model dengan tingkat akurasi yang tinggi, yang didasarkan pada hasil prediksi data evaluasi dari masing-masing penelitian. Dalam penelitian yang dilakukan oleh Zhou et al. (2021), dijelaskan bahwa model LSTM memiliki keterbatasan dalam menangani data *non-linear*. Oleh karena itu, sebuah model gabungan yang menggunakan RNNs dan LSTM dikembangkan untuk menghasilkan model yang memiliki kemampuan *Long-Term dependencies* yang lebih baik. Diharapkan bahwa model serupa akan mampu mengatasi data dengan rentang waktu yang berbeda dalam penelitian yang akan dilakukan.

## Rumusan Masalah

Berdasarkan penjelasan yang telah disampaikan pada latar belakang, Secara umum, Aset Integrity menjadi perhatian utama dalam keberlangsungan operasional industri minyak dan gas. Kegagalan pada Asset Integrity dapat menyebabkan resiko kerusakan lingkungan dan korban jiwa, selain itu sistem tradisional dalam penanganan Asset Integrity membutuhkan waktu dan biaya yang mahal dan cenderung tidak optimal. Oleh karena itu, diperlukan sebuah penelitian yang memanfaatkan kemampuan Intelligent Asset Integrity untuk merekomendasikan tindakan pengamanan, perbaikan atau pemeliharaan, mengurangi risiko kegagalan, dan memaksimalkan efisiensi operasional. Berdasarkan latar belakang yang telah disebutkan, diperoleh pertanyaan peneltitian sebagai berikut:

1. Bagaimana meningkatkan performa model standar machine learning untuk memprediksi kinerja Asset Integrity di industri minyak dan gas ?
2. Bagaimana hasil evaluasi model machine learning pada Intelligent Asset Integrity di industri minyak dan gas ?

## Batasan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, diperoleh batasan masalah yang dalam penelitian ini adalah:

1. Data yang digunakan berasal dari Petronas
2. Data yang digunakan merupakan suhu aset per satuan waktu (time series)
3. Rentang data yang digunakan dari tahun 2020 sampai 2023
4. Metodologi dengan pendekatan CRISP-DM dibatasi hingga tahap Evaluation

## Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, maka tujuan penelitian yang ingin dicapai adalah:

1. Untuk meningkatkan performa model standar machine learning untuk memprediksi kinerja Asset Integrity di industri minyak dan gas.
2. Untuk mengevaluasi efektivitas penggunaan model machine learning pada Asset Integrity di industri minyak dan gas.

## Manfaat Penelitian

Brdasarkan latar belakang yang telah diuraikan di atas, diperoleh manfaat perancangan yang ingin dicapai adalah:

1. Manfaat Teoritis

Sebagai sarana untuk menambah ilmu pengetahuan di bidang machine learning, Asset Integrity, serta industri minyak dan gas.

1. Manfaat Praktis

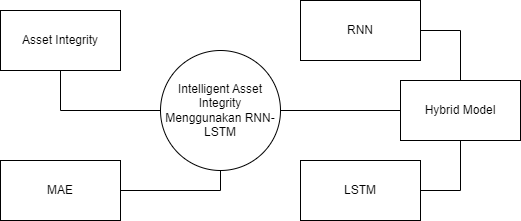
Hasil penelitian diharapkan dapat memberikan rekomendasi memaksimalkan efisiensi operasional dan keamanan pada industri minyak dan gas.

A person standing in a tree

Description automatically generated

# BAB II TINAJAUAN PUSTAKA

## Research Puzzle



Gambar 2.1 Research Puzzle

## Tinjauan Pustaka

Penelitian mengenai Intelligent Aset Integrity telah banyak dilakukan, terutama di industri minyak dan gas. Penelitian tersebut tersebar pada beberapa orientasi yang dicapai, Sun et al. (2019) dan Yang et al. (2020) pada penelitiannya membuat model untuk mendeteksi kerusakan pada aset. Al-maythony et al. (2019), Sun et al. (2020), dan Zhou et al. (2021) melakukan penelitian untuk memperoleh Intelligent Aset Integrity yang dapat melakukan prediksi sesuai data aset yang dipelajari. Selain itu, Aset Integrity dapat dipengaruhi oleh beberapa variabel, baik berkaitan secara eksternal maupun internal. Setiawan et al. (2021) dan Tang et al. (2019) pada penelitiannya menunjukkam bahwa Intelligent Asset Integrity dapat menentukan variabel signifikan dan potensial yang mempengaruhi Asset Integrity suatu aset.

Metode yang diterapkan setiap penelitian memiliki kelemahannya masing-masing, data yang dimiliki menjadi pertimbangan dalam menentukan metode yang tepat untuk membangun model. Penelitian Yang et al. (2018) memanfaatkan kelebihan RNN yang mampu mengatasi data dengan pola perubahan terhadap waktu, sehingga mampu melakukan prediksi terhadap sisa masa pakai mesin pesawat. Selain itu, penelitian Li et al. (2020) dan Wang et al. (2019) menunjukkan dua buah data dengan hubungan non-linear dapat diselesaikan dengan baik menggunakan RNN. Namun, pada penelitian berbeda RNN memiliki kelemahan dalam mengatasi data dengan pola ketergantungan jangka panjang. Hal ini telah ditunjukkan pada penelitian Yang Et al. (2017) dan Wang et al. (2018) dimana model yang dihasilkan memiliki akurasi yang rendah pada data kasus hasil prediksi yang bergantung pada pola jangka panjang.

Kelemahan model menggunakan arsitektur RNN dapat diatasi dengan cara menutupi keterbatasan model melalui penggabungan model dengan arsitektur berbeda. Model dengan kelebihan pada sisi yang menjadi kelemahan RNN diperhitungkan sebagai pasangan arsitektur yang cocok dalam pembuatan hybrid model. Penelitian Zhou et al. (2021) memanfaatkan penggabungan arsitektur RNN dan LSTM untuk memprediksi masa sisa pakai turbin angin, LSTM memiliki kemampuan yang baik dalam memprediksi data dengan ketergantungan jangka panjang (Long Term) (Sakinah et al, 2019). Hasil dari penelitian Zhou et al. (2021) menunjukkan bahwa hybrid model tersebut mampu mengatasi kelemahan arsitektur RNN, model mampu mengatasi ketergantungan jangka panjang dan memberikan performa yang baik pada hasil evaluasinya.

Random Forest, CNN

## Landasan Teori

* + 1. **RNN**

Recurrent Neural Network (RNN) merupakan sebuah model dari jaringan saraf tiruan yang cocok digunakan untuk klasifikasi pola dengan masukan dan keluaran sistem berupa data sequensial. Arsitektur RNN biasanya memiliki kelemahan ketika mempelajari informasi dengan interval yang panjang (long term dependencies) (Alghifari et al 2022). RNN adalah sejenis jaringan saraf tiruan yang memiliki kemampuan untuk melihat korelasi tersembunyi yang terjadi pada data (Tian et al, 2018).

RNN dapat melakukan proses encoding pada informasi sebelumnya dalam proses pembelajaran yang memiliki lapisan tersembunyi. Sehingga, data dari *time series* dapat dipelajari secara efektif (Xu et al. 2018). RNN cocok untuk data time series karena memiliki kemampuan memahami pola ketergantungan jangka pendek selama periode variabel (Mahmoud dan Mohammed, 2021). RNN termasuk dalam kelas jaringan saraf yang dapat memproses data sekuensial. RNN menerima rangkaian vektor (x1,x2,..,xn) sebagai masukan dan menghasilkan rangkaian kedua (h1,h2,..,hn) yang mengandung informasi tentang urutan masukan pada setiap Langkah (Rathore et al, 2018). Persamaan transisi RNN adalah sebagai berikut (Rathore et al, 2018) :

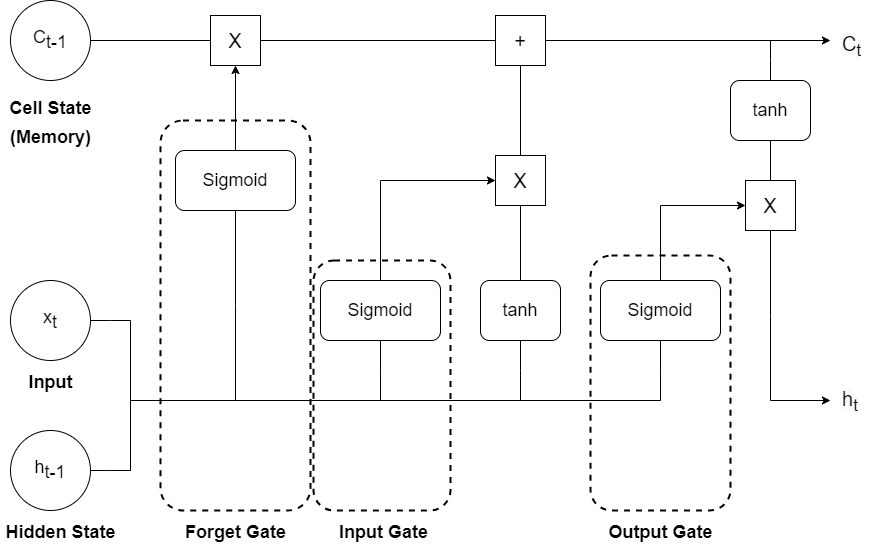
(2.1)

(2.2)

Dalam konteks ini, fungsi g merujuk pada fungsi terbatas dan halus seperti tangen hiperbolik atau fungsi logistik. merupakan nilai output atau prediksi dari jaringan saraf berulang. Vektor input jaringan , bersama dengan status tersembunyi sebelumnya dan bias b, menentukan status tersembunyi berulang pada setiap langkah waktu t. Matriks bobot, W dan U, berfungsi sebagai filter yang menentukan seberapa signifikan input dan status tersembunyi dari masa lalu yang akan diberikan.

* + 1. **LSTM**

LSTM merupakan salah satu jenis Recurrent Neural Network. LSTM yang memiliki kemampuan mempelajari data di mana dalam setiap neuron LSTM memiliki beberapa gerbang yang mengatur memori dari setiap neuron itu sendiri (Sakinah et al., 2019). LSTM adalah metode yang dapat digunakan untuk mempelajari suatu pola pada data deret waktu (Ningrum et al, 2021). Pada penelitian Fischer dan Kauss (2018), LSTM network menunjukkan keunggulan model prediksi yang lebih baik dari pada model-model standar lainnya, seperti : random forest, a standard deep neural network and a standard logistic regression. Berikut arsitektur dari LSTM :



Gambar 2.2 Arsitektur Jaringan LSTM (Rathore et al, 2018)

Persamaan transisi LSTM adalah sebagai berikut (Rathore et al, 2018) :

(2.3)

(2.4)

(2.5)

(2.6)

(2.7)

(2.8)

(2.9)

Dalam Persamaan (3) hingga (9),, , dan masing-masing merujuk pada input gate, forget gate, dan output gate. Selain itu, menunjukkan sel memori, fungsi aktivasi diberikan oleh , status tersembunyi LSTM ditunjukkan oleh , dan nilai prediksi diberikan oleh . Berbeda dengan unit recurrent neural network, di mana cakupannya hanya diubah setiap langkah waktu, LSTM memiliki kemampuan untuk memutuskan apakah akan mempertahankan memori saat ini dengan bantuan gerbang tambahan.

* + 1. **Asset Integrity**

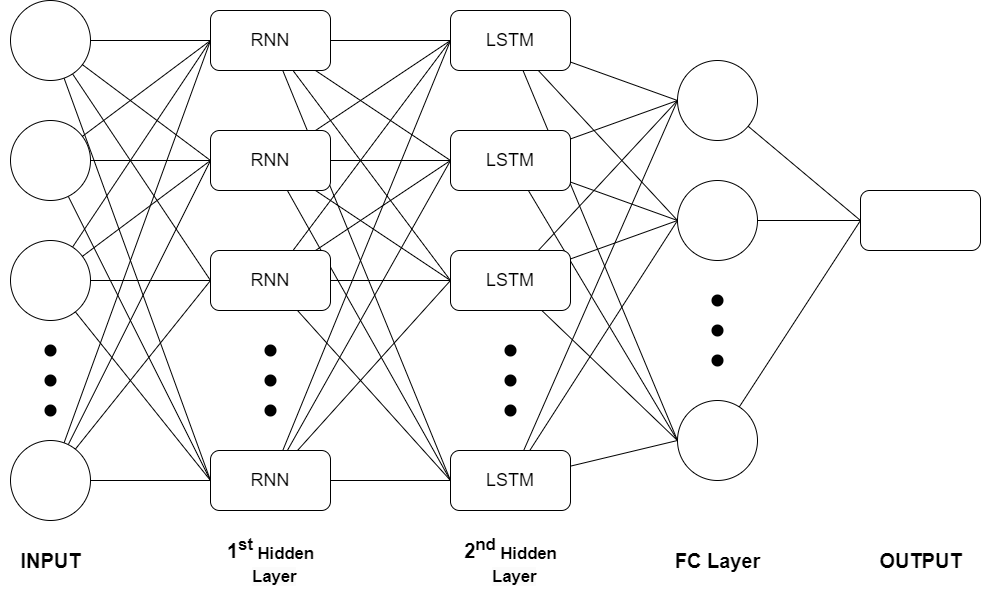
Asset Integrity berada di pusat manajemen keselamatan dan memiliki fitur yang tumpang tindih dengan keselamatan proses. Asset Integrity mencakup pengelolaan orang, sistem, proses, dan sumber daya untuk memastikan aset beroperasi dengan risiko minimal bagi karyawan, masyarakat, dan lingkungan (Tang et al, 2021). Keputusan Asset Integrity untuk peralatan yang kritis dalam produksi sebagian besar bersifat kualitatif dan didorong oleh pengalaman (Reza, 2020).

* + 1. **Hybrid Model**

Hybrid Model adalah sekelompok algoritme sederhana yang bekerja bersama untuk melengkapi dan meningkatkan satu sama lain. Dengan bekerja bersama, mereka dapat menyelesaikan masalah yang tidak dapat mereka selesaikan sendiri. Dalam Pembelajaran Model Hibrida (Hybrid Model Learning), terdapat berbagai jenis teknik yang berinteraksi dengan data dan cara yang berbeda (Domo, 2023). Model hibrida digunakan dalam pembuatan model prediksi untuk mengatasi kelemahan salah satu model (Atlan et al., 2021).

Dalam membangun hybrid model, proses penggabungan dua model yang berbeda menjadi satu disebut sebagai Ensemble Learning. Van et al (2020) dalam jurnalnya membahas pembuatan model prediksi dengan metode LSTM dan metode Elman, dilakukan Ensemble learning sehingga diperoleh hybrid network. Metode Elman pada penelitian ini merupakan solusi dari keterbatasan metode LSTM dengan kemampuan short-term memory. Wu et al (2019) membuktikan bahwa Ensemble Learning merupakan metode yang efektif dalam melakukan prediksi pergerakan harga minyak.

Pada penelitian ini, model hybrid terdiri dari gabungan antara RNN dan LSTM. Islam et al (2020) menjelaskan bahwa kedua model jaringan saraf ini adalah teknologi canggih yang dapat mengungguli prediksi berbasis regresi dalam hal akurasi. Berdasarkan studi yang dilakukan oleh Faru et al (2023), dalam penggabungan ini, hybrid model tersebut dibangun dengan layer LSTM yang ditempatkan sebelum layer RNN. Berikut adalah arsitektur hybrid model RNN-LSTM:



Gambar 2.3 Arsitektur Model Hybrid RNN-LSTM (Rathore et al, 2018)

* + 1. **MAE**

Pada saat melakukan evaluasi, Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE) secara luas digunakan dalam banyak sistem rekomendasi untuk mengukur perbedaan antara skor yang diprediksi dan penilaian aktual pengguna, seperti metode Slope One (Wang dan Lu, 2023). Dalam konteks Machine Learning mean absolute error merupakan besarnya perbedaan antara prediksi suatu observasi dengan nilai sebenarnya dari observasi tersebut. MAE mengambil rata-rata *absolute error* untuk sekelompok prediksi dan observasi sebagai pengukuran besarnya kesalahan untuk seluruh kelompok tersebut (Rustam et al, 2020).

* + 1. **Optimisasi**

Stochastic Gradient Descent (SGD) adalah salah satu algoritma optimisasi yang banyak digunakan dalam berbagai jenis model machine learning, termasuk model time series. Dalam SGD, optimisasi dilakukan dengan cara menghitung gradien (turunan parsial) dari fungsi kerugian (*loss* function) terhadap setiap parameter-model. Gradien ini menunjukkan arah dan besaran langkah yang harus diambil untuk mengurangi nilai fungsi kerugian tersebut. Proses ini dilakukan secara iteratif untuk memperbarui parameter-model hingga mencapai titik konvergensi atau nilai fungsi kerugian yang sudah mencukupi (Tian et al, 2023).

Fungsi aktivasi ReLU (Rectified Linear Unit) sering digunakan untuk mengatasi masalah vanishing gradient pada jaringan saraf yang dalam. Masalah vanishing gradient terjadi ketika gradien menjadi sangat kecil saat menyebar kembali melalui banyak lapisan selama proses pelatihan (Tan et al, 2019).

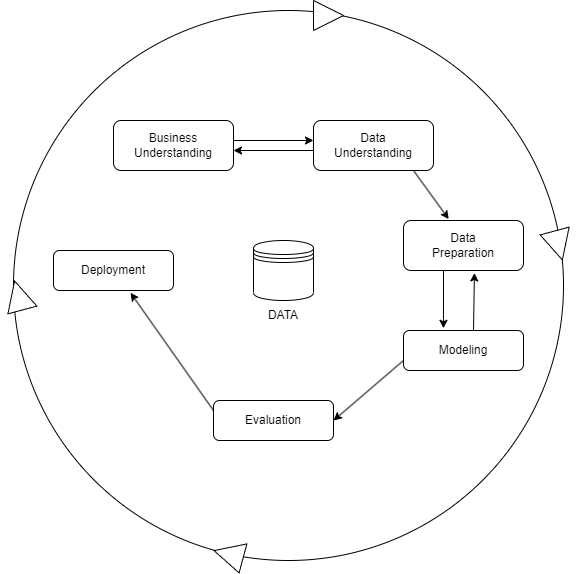
Attention layer adalah komponen kunci dalam model neural network yang memungkinkan model untuk fokus pada bagian-bagian penting dari input yang diberikan, sehingga meningkatkan kemampuan model untuk mengatasi ketergantungan jarak jauh dalam data, seperti pada tugas-tugas pemrosesan bahasa alami dan pemodelan urutan, termasuk time series (Sharaf Al-deen et al, 2021). Pada pemodelan machine learning yang dilakukan Dandala et al, (2019), arsitektur attention layer diletakkan setelah hidden layer dan berada sebelum layer output. Hal serupa juga dilakukan oleh Yin et al, (2022) dalam penelitiannya dimana attention layer diletakkan sebelum layer terakhir (output).

A person standing in a tree

Description automatically generated

# BAB III METODE PENELITIAN

Metodologi yang digunakan menggunakan pada tahapan penelitian ini menggunakan pendekatan CRISP-DM, seperti berikut:



Gambar 3.1 Proses CRISP-DM (Schröer et al, 2021)



## Bentuk Penelitian

Penelitian kuantitatif adalah penelitian yang dimaksud untuk mengungkapkan gejala secara holistik-konstektual melalui pengumpulan data dari latar alami dengan memanfaatkan diri peneliti sebagai instrumen kunci. Penelitian kuantitatif menitikberatkan pada pengukuran dan hubungan sebab-akibat antara bermacam macam variabel, bukan prosesnya, penyelidikan dipandang berada dalam kerangka bebas nilai. (Syahrum dan Salim, 2014).

## Pengumpulan Data

Kumpulan data yang digunakan dalam penelitian ini diperoleh dari perusahaan industri minyak dan gas Petronas.

## Tahapan Penelitian

Berikut tahapan penelitian yang dilakukan pada penelitian ini:



### Business Understanding

Asset Integrity mengacu pada kondisi dan keandalan suatu aset. Keberlanjutan operasional dan keamanan personel sangat bergantung pada integritas yang dipertahankan oleh aset-aset tersebut. Namun, terdapat beberapa masalah yang perlu diidentifikasi terkait Asset Integrity ini. Pertama, kerusakan atau kegagalan aset dapat mengakibatkan konsekuensi yang serius seperti kecelakaan, gangguan operasional, penurunan produksi, dan dampak lingkungan yang merugikan. Terutama dalam industri minyak dan gas, di mana aset sering beroperasi dalam kondisi ekstrem, seperti tekanan tinggi, suhu ekstrim, dan lingkungan korosif, risiko kerusakan seperti korosi, kelelahan material, retak, dan penurunan kinerja aset semakin meningkat seiring waktu.

### Data Understanding

Beberapa hal yang akan menjadi fokus perhatian dalam memahami karakteristik dari data yang diperoleh adalah sebagai berikut:

1. Memiliki ketergantungan terhadap waktu, nilai suhu pada waktu tertentu hanya dipengaruhi oleh nilai-nilai sebelumnya dalam seri waktu tersebut, dan tidak dipengaruhi oleh faktor eksternal atau variabel lain di luar waktu itu sendiri.
2. Asumsi data yang digunakan memiliki rata-rata, varians, dan kovarians suhu tetap konstan seiring waktu. Dalam konteks suhu aset, ini berarti bahwa suhu aset secara keseluruhan tidak mengalami perubahan sistematis dalam jangka waktu tertentu.
3. Tidak ada pola musiman, tidak ada pola yang terlihat dalam data. Dalam konteks suhu aset, ini berarti bahwa tidak ada fluktuasi periodik yang teratur dalam suhu aset terkait dengan musim, waktu harian, atau interval waktu lainnya.

### Data Preparation

Pada saat persiapan data terdapat beberapa hal yang penting untuk dilakukan seperti pembersihan data, integrasi atau penggabungan data, transformasi data. Pada tahap pembersihan data, dilakukan identifikasi dan penanganan nilai yang hilang, duplikat, atau outlier dalam data. Duplikat dapat dihapus atau diabaikan, dan outlier dapat diidentifikasi dan diperlakukan sesuai kebijakan analisis yang ditetapkan. Selanjutnya, integrasi data melibatkan penggabungan data dari berbagai sumber menjadi satu dataset yang konsisten. Hal ini melibatkan pemadanan skema, penanganan konflik data, atau penggabungan data berdasarkan atribut yang relevan. Setelah itu, transformasi data dilakukan untuk mengubah atau mengubah format data agar sesuai dengan kebutuhan analisis.

### Modeling

Setelah mempelajari data yang dimiliki dan melakukan asumsi terhadap data dengan mempertimbangkan batasan pembuatan model, maka langkah berikutnya adalah membangun model LSTM dan RNNs secara hybrid (gabungan) sesuai pedoman literatur dan penelitian sebelumnya. Kemudian, model akan dilatih dengan data yang sudah disiapkan sehingga dapat dilakukan prediksi pada komponen Asset Integrity.

### Evaluation

Setelah memperoleh model prediksi, penting untung melakukan evaluasi. Evaluasi model merupakan tahap penting untuk memahami kinerja dan kehandalan model tersebut. Evaluasi akan dilakukan dengan menggunakan metrik evaluasi untuk mengukur kinerja model prediksi. Kemudian membuat grafik korelasi antara hasil prediksi dan hasil sebenarnya, hal ini bertujuan untuk memperoleh kemudahan visualisasi pada hasil evaluasi.

### Deployment

Tahapan terakhir adalah melakukan pengujian pada model yang sudah dibuat, dimana hasil dari pemodelan akan digunakan untuk melakukan prediksi secara langsung pada data yang telah diperoleh. Hasil prediksi kemudian akan digunakan sebagai wawasan untuk mendukung operasional, keamanan, dan pengoptimalan kerja yang berbasis manual. Hal ini berdasarkan dari hasil prediksi yang diberikan oleh model.

A person standing in a tree

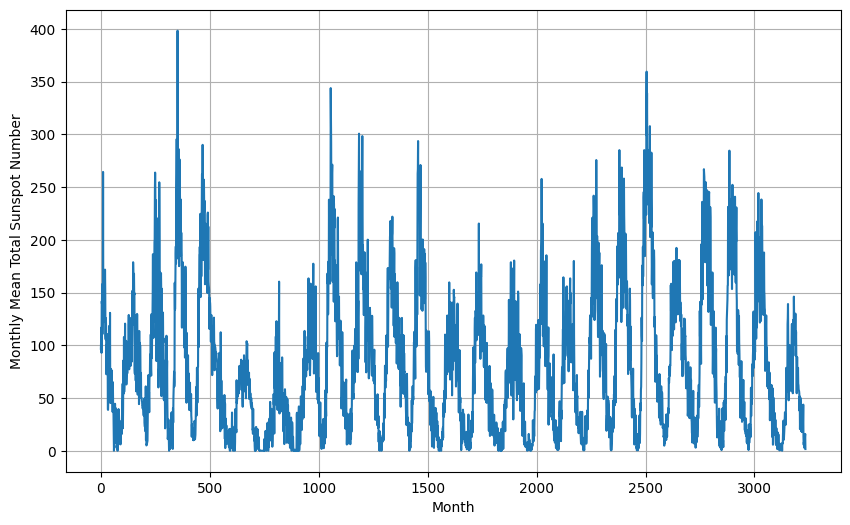
Description automatically generated

# BAB IV HASIL DAN PEMBAHASAN



## Data

Kumpulan data yang digunakan dalam penelitian ini diperoleh dari perusahaan industri minyak dan gas Petronas. Namun, dalam pembangunan model dibutuhkan data yang lengkap dan banyak. Hal ini untuk memperoleh cakupan yang lebih akurat mengenai model yang sedang dibangun. Berikut visualisasi data time series (independent) yang digunakan dalam penelitian ini:



Gambar 4.1 Grafik Data Independen (Sunspot)

Data ini memiliki nilai penting dalam pemodelan time series karena data tersebut mempertahankan struktur waktu yang mengikuti urutan pengamatan. Ini memungkinkan pengenalan pola, tren, dan fluktuasi periodik yang terjadi dalam jangka waktu tertentu. Data ini dapat digunakan untuk melakukan peramalan tentang aktivitas bintik matahari di masa depan sehingga dapat membantu dalam mendeteksi anomali atau peristiwa luar biasa yang mungkin terjadi pada aktivitas bintik matahari. Pemantauan terhadap peristiwa seperti ledakan matahari dan aktivitas luar biasa lainnya memiliki relevansi yang signifikan terhadap data time series pada umumnya.

## Preprocessing Data

### Pembersihan Data

Pembersihan data merupakan tahap penting dalam analisis data yang dilakukan untuk mengatasi ketidakkonsistenan dan ketidaklengkapan nilai dalam dataset. Langkah pertama dalam proses ini adalah mengidentifikasi variabel yang memiliki data yang hilang (missing data), di mana perlu diperiksa apakah nilai yang tidak ada tersebut dapat diisi dengan nilai lain atau harus dihapus. Jika mungkin, teknik imputasi seperti penggunaan rata-rata, median, modus, atau interpolasi dapat digunakan untuk mengisi nilai yang hilang. Namun, dalam situasi tertentu, jika data yang hilang sangat signifikan atau mempengaruhi hasil analisis secara substansial, maka penghapusan data tersebut mungkin lebih sesuai.

### Split Data

Untuk membagi dataset menjadi bagian pelatihan dan pengujian, penelitian ini menggunakan perbandingan 80% dan 20%. Ini berarti bahwa 80% data akan digunakan untuk melatih model, sementara 20% sisanya akan digunakan untuk menguji kinerja model. Pembagian ini penting untuk menghindari overfitting, di mana model terlalu mempelajari data pelatihan dan tidak dapat mengeneralisasi dengan baik pada data baru. Dengan menggunakan 80% data pelatihan, penelitian ini memberikan kesempatan yang cukup bagi model untuk mempelajari pola dan fitur yang ada dalam dataset. Sementara itu, 20% data pengujian memberikan evaluasi obyektif tentang sejauh mana model dapat menghasilkan prediksi yang akurat pada data yang belum pernah dilihat sebelumnya.

### Feature dan Labeling Data

Pemberian label dan feature pada dataset menggunakan Teknik window dataset. Window dataset adalah teknik pengolahan data di mana data asli dibagi menjadi jendela-jendela kecil atau segmen-segmen yang saling tumpang tindih. Setiap jendela ini berisi sejumlah data yang saling terkait. Teknik ini umumnya digunakan dalam konteks pemodelan deret waktu atau analisis sinyal untuk menangkap pola atau tren dalam data yang berkelanjutan. Dalam konteks ini, kita ingin memprediksi atau mengklasifikasikan sesuatu berdasarkan pola atau tren yang ada dalam data sebelumnya.

## Pemodelan dengan LSTM-RNN

Berikut model RNN-LSTM yang digunakan:

A screenshot of a computer

Description automatically generated

Model tersebut merupakan model gabungan dari (Long Short-Term Memory) dan RNN (Recurrent Neural Network) yang dibangun menggunakan library TensorFlow. Pertama, model ini menggunakan Conv1D layer untuk melakukan operasi konvolusi satu dimensi pada input. Layer ini memiliki 64 filter dengan ukuran kernel 3, menggunakan fungsi aktivasi ReLU, padding tipe 'causal', dan menerima input dengan bentuk [window\_size, 1], window\_size menunjukkan ukuran jendela untuk input. Tujuan Conv1D layer adalah untuk mengekstraksi fitur-fitur penting dari data input.

Selanjutnya, model ini menggunakan SimpleRNN layer dengan 64 unit dan return\_sequences=True. Layer ini digunakan untuk mempertahankan urutan output lengkap dari SimpleRNN layer dan mengirimkannya sebagai input ke layer berikutnya. Selanjutnya, model ini memiliki layer LSTM dengan 64 unit. LSTM (Long Short-Term Memory) adalah jenis arsitektur jaringan saraf rekuren (RNN) yang dirancang khusus untuk menangani masalah gradien yang hilang pada RNN konvensional. Layer LSTM ini membantu model untuk menangkap pola temporal yang kompleks dalam data.

Setelah itu, model ini memiliki Dense layer dengan satu unit. Layer ini bertanggung jawab untuk menghasilkan output akhir yang diinginkan oleh model. Terakhir, model ini menggunakan Lambda layer yang menerapkan fungsi lambda untuk mengalikan output model dengan 100. Tujuan dari langkah ini adalah untuk mengubah skala output agar sesuai dengan rentang nilai yang diharapkan.

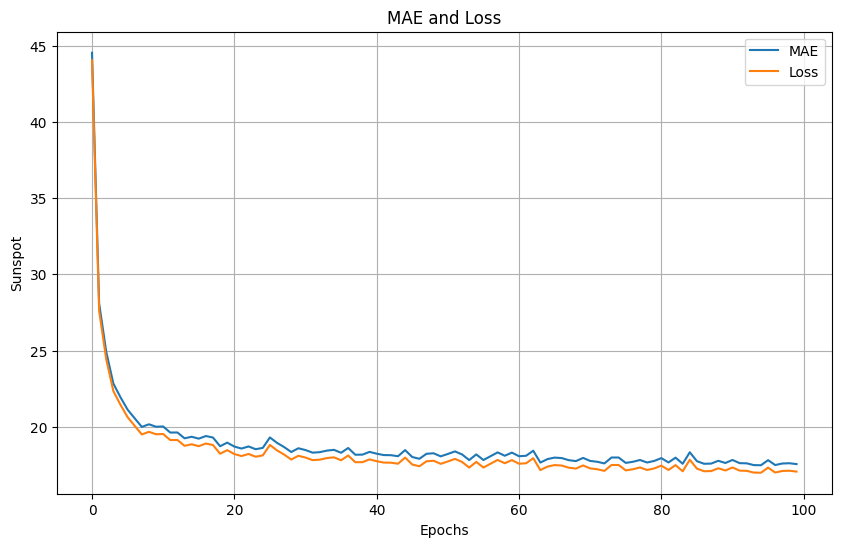
## Matriks Evaluasi (MAE & *LOSS*)

Matriks Evaluasi dibangun menggunakan Mean Absolute Error dari proses evaluasi data train dan data evaluation. Hal ini bertujuan, untuk membantu visualisasi evaluasi model dengan menggunakan data yang tercatat dalam objek *history*. Pada awal fungsi, MAE dan *loss* diambil dari log *history*. Variabel MAE akan berisi nilai mean absolute error (MAE) yang tercatat dalam *history*, sedangkan variabel *loss* akan berisi nilai *loss* yang tercatat dalam *history*.

Selanjutnya, variabel epochs diinisialisasi sebagai range yang berdasarkan panjang (length) dari variabel *loss*. Hal ini dilakukan untuk mendapatkan jumlah epoch yang sesuai dengan jumlah data yang tercatat dalam log *history*. Dengan memanfaatkan informasi yang diperoleh dari *history*, fungsi tersebut dapat digunakan untuk memvisualisasikan perubahan MAE dan *loss* sepanjang epoch. Visualisasi ini dapat memberikan wawasan mengenai performa model dan perkembangannya selama proses pelatihan.

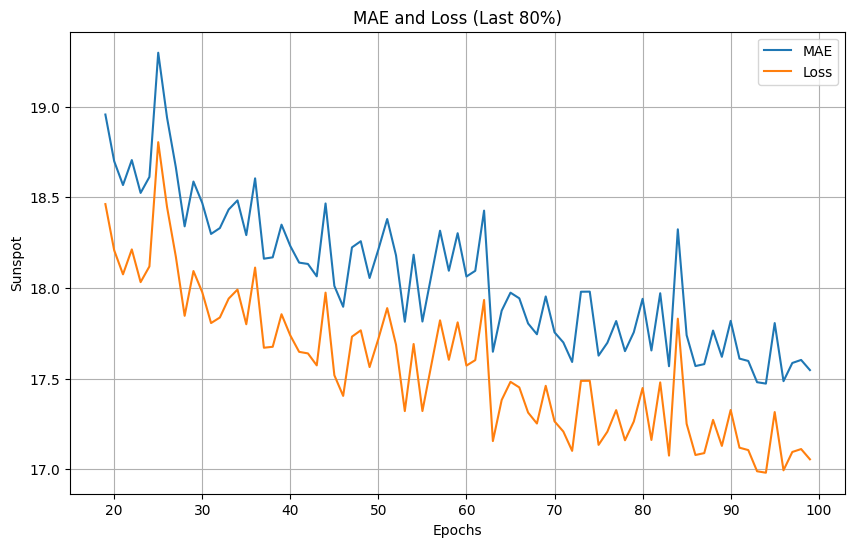
### Hasil Evaluasi LSTM-RNN

Berikut hasil evaluasi menggunakan model LSTM-RNN berdasarkan *history* *loss* dan MAE yang dihasilkan selama proses training:



Gambar 4.8 Grafik Evaluasi Model LSTM-RNN

Dalam proses analisis evaluasi model LSTM-RNN, terdapat manfaat dalam memisahkan grafik awal pelatihan di mana terjadi penurunan nilai yang signifikan. Hal ini dilakukan untuk memperoleh wawasan yang lebih baik tentang perbandingan proporsi antara MAE dan *loss* serta tingkat kesalahan dan keakuratan model. Oleh karena itu, berikut ini disajikan plot 80% terakhir dari grafik evaluasi tersebut:



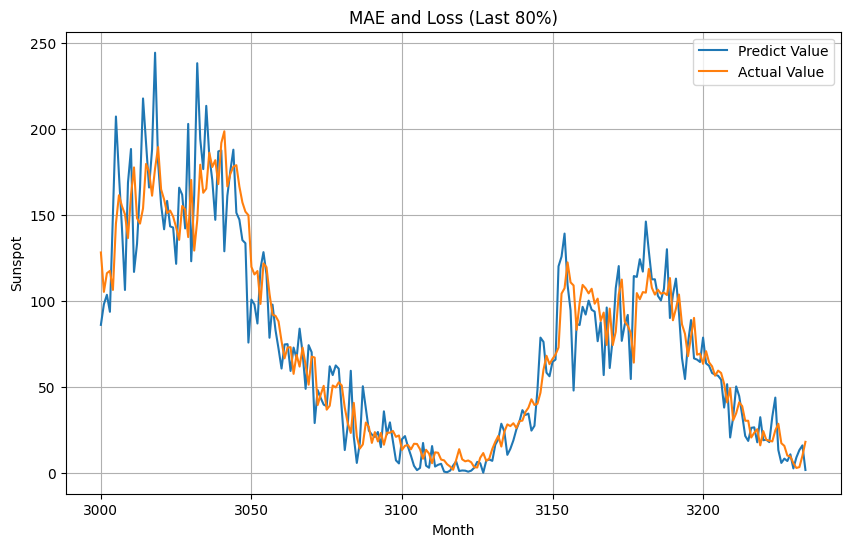
Gambar 4.9 Grafik Evaluasi Model LSTM-RNN (80%)

Grafik *history* *loss* dan MAE pada model LSTM-RNN tersebut menunjukkan penurunan yang signifikan dalam Mean Absolute Error (MAE) dan *loss* pada awal pelatihan serta kecenderungan penurunan yang berlanjut hingga akhir epoch menandakan adanya kemajuan yang baik dalam proses pembelajaran model. Hasil ini menggambarkan bahwa model mampu menyesuaikan diri dengan data pelatihan dan secara efektif mengurangi kesalahan prediksi. Ketika MAE dan *loss* memiliki selisih yang tidak terlalu besar, hal ini menunjukkan bahwa model tersebut tidak mengalami overfitting secara berlebihan.

## Hasil Prediksi (MAE)

Proses prediksi dilakukan dengan beberapa langkah. Pertama, data waktu yang akan diprediksi diambil dari data asli dengan memperhatikan jendela waktu yang ditentukan. Selanjutnya, prediksi dilakukan menggunakan model yang sudah disiapkan. Hasil prediksi kemudian diproses dan dibandingkan dengan data validasi. Hasil prediksi dan data validasi kemudian ditampilkan dalam sebuah plot. Selain itu, juga dihitung metrik MAE (Mean Absolute Error) untuk mengevaluasi seberapa akurat prediksi model.

Berikut berturut-turut hasil prediksi dari model LSTM-RNN menggunakan Mean Absolute Error (MAE):



Gambar 4.13 Grafik Prediksi Model LSTM-RNN

Hasil prediksi time series pada data test menggunakan model LSTM-RNN tersebut menunjukkan nilai yang tidak terlalu jauh berbeda dengan nilai sebenarnya menandakan bahwa model tersebut memiliki tingkat konsistensi yang baik dalam melakukan prediksi Grafik tersebut tidak menunjukkan terdapat momen saat nilai error yang cukup siginifkan antara nilai sebenarnya dan hasil prediksi. Rata-rata hasil dari keseluruhan hasil prediksi tersebut adalah 14.826439 (MAE)

## Conifdence Interval

Confidence interval (CI) adalah estimasi interval statistik yang digunakan untuk memberikan batasan kisaran nilai yang mungkin mengandung parameter populasi dengan tingkat keyakinan tertentu. CI didasarkan pada analisis sampel data dan digunakan untuk memberikan informasi tentang ketidakpastian dalam estimasi parameter populasi, seperti rata-rata, proporsi, atau variansi.

Confidence menggambarkan sejauh mana ketidakpastian dalam estimasi model statistik, seperti parameter model atau prediksi model. Pada model yang menggunakan SGD (Stochastic Gradient Descent), confidence interval dapat membantu dalam memahami sejauh mana nilai-nilai parameter model dapat bervariasi. Studi ini melibatkan serangkaian langkah yang terstruktur untuk mengevaluasi model yang telah dirancang. Untuk mengukur ketidakpastian dalam estimasi model, langkah pertama yang diambil adalah menjalankan model sebanyak 100 kali dengan 100 epoch pada setiap percobaan. Penting untuk dicatat bahwa data yang digunakan dalam setiap iterasi tetap konsisten, sehingga fokus utama penilaian adalah pada parameter model yang telah dihasilkan.

Selanjutnya, data yang diperoleh dari menjalankan model sebanyak 100 kali serta melakukan evaluasi dan prediksi pada data evaluasi, hasil tersebut diproses dengan menggunakan confidence interval dengan tingkat kepercayaan 95%. Proses ini melibatkan pemotongan 2.5% dari batas atas dan 2.5% dari batas bawah data. Dengan demikian, kita dapat memperoleh perkiraan kisaran atau rentang yang mungkin untuk parameter model. Hasil analisis akan divisualisasikan untuk mempermudah interpretasi. Rata-rata dari mean absolute error (MAE), hasil prediksi, dan *loss* pada model akan disajikan dalam bentuk grafik atau diagram. Hal ini akan membantu dalam memahami performa model secara keseluruhan dan memberikan gambaran tentang tingkat ketidakpastian yang terkait dengan parameter model.

### Confidence Interval LSTM-RNN

Berikut hasil evaluasi menggunakan model LSTM-RNN berdasarkan *history* *loss*, MAE, dan hasil prediksi pada 100 kali percobaan yang dihasilkan selama proses training:

A graph of a graph

Description automatically generated

Gambar 4.26 Grafik Confidence Interval MAE pada model LSTM-RNN

A graph of a loss

Description automatically generated

Gambar 4.27 Grafik Confidence Interval *LOSS* pada model LSTM-RNN

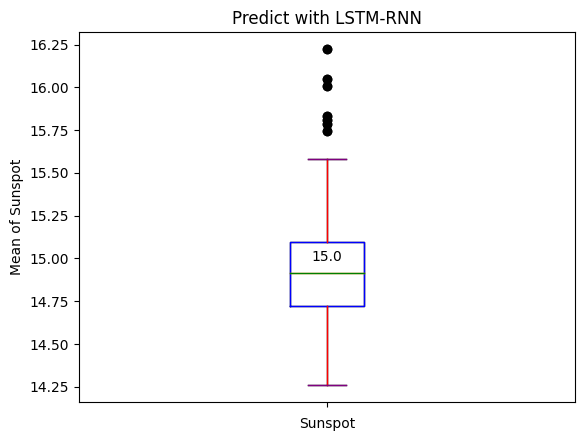
Grafik 4.26 dan 4.27 tersebut menampilkan perjalanan dari *history* *loss* dan MAE model RNN-LSTM selama 100 kali iterasi. Garis biru menggambarkan nilai terbesar dari *history* *loss* dan MAE, yang tercapai pada suatu titik selama iterasi. Garis hijau merepresentasikan nilai terkecil dari *history* *loss* dan MAE, yang mencerminkan performa terbaik model atau analisis pada iterasi tertentu. Selanjutnya, rata-rata dari *history* *loss* dan MAE dihitung dan direpresentasikan sebagai garis merah. Dari grafik tersebut juga terlihat bahwa *history* *loss* secara konsisten menurun seiring dengan berjalan-nya iterasi. Hal ini menunjukkan bahwa model atau analisis yang digunakan secara umum semakin mendekati hasil yang lebih akurat atau optimal seiring berjalan-nya waktu atau iterasi.

A graph with red and blue lines

Description automatically generated

Gambar 4.28 Grafik Confidence Interval hasil prediksi pada model LSTM-RNN

Grafik 4.28 tersebut menggambarkan hasil prediksi (garis biru dan garis hijau) dari model LSTM-RNN ketika dijalankan 100 kali. Garis biru mewakili nilai terbesar yang dihasilkan dari prediksi tersebut, sedangkan garis hijau merepresentasikan nilai terkecil dari hasil prediksi. Selanjutnya, hasil prediksi yang diperoleh dari model atau analisis tersebut dicocokkan dengan nilai sebenarnya dan direpresentasikan sebagai garis merah. Berdasarkan visualisasi tersebut, tampak bahwa hasil prediksi (garis biru dan garis hijau) relatif konsisten dan berada di sekitar garis merah (hasil aktual). Hal ini menunjukkan bahwa model atau analisis tersebut memberikan prediksi yang relatif mendekati nilai aktual dari data, dan memiliki tingkat kesalahan yang cukup kecil.



Gambar 4.29 Distribusi MAE hasil prediksi pada model LSTM-RNN

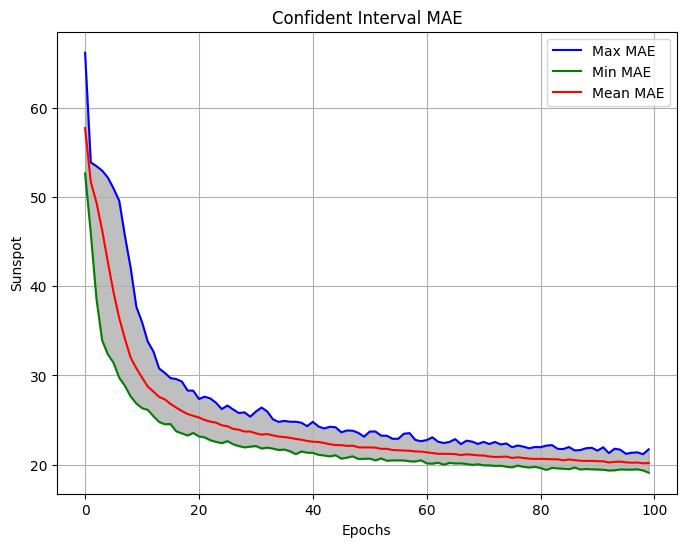
Distribusi Mean Absolute Error (MAE) dari hasil prediksi dengan model LSTM-RNN atau ditampilkan menggunakan box plot. Kotak pada box plot ini merepresentasikan kuartil pertama (Q1) hingga kuartil ketiga (Q3) dari distribusi MAE. Garis tengah kotak menunjukkan nilai median dari distribusi. Selain itu, garis di luar kotak menandai rentang data, dimana titik merepresentasikan adanya outlier (nilai ekstrem). Rata-rata dari distribusi MAE ini adalah 15.0, yang menggambarkan tingkat kesalahan prediksi secara keseluruhan. Melalui visualisasi ini, dapat dilihat sebaran dan kecenderungan kesalahan prediksi dari model yang digunakan.

Dengan menggunakan pendekatan ini, penelitian ini akan memberikan wawasan yang lebih dalam tentang bagaimana model berperilaku dalam berbagai iterasi dan bagaimana ketidakpastian dalam parameter dapat mempengaruhi hasil prediksi. Dengan demikian, langkah-langkah ini membentuk pendekatan ilmiah yang ketat untuk menguji dan mengevaluasi model yang dikembangkan dalam penelitian ini.

## Attention Layer

Berdasarkan literatur review, penggunaan attention layer pada model neural network, terutama pada data time series, telah terbukti efektif dalam meningkatkan kemampuan dan kehandalan model. Attention layer memungkinkan model untuk lebih fokus pada bagian penting dari data input dan memberikan bobot yang lebih tinggi pada informasi yang lebih relevan dalam setiap langkah waktu. Pada tahap sebelumnya, model yang telah memberikan hasil prediksi terbaik digunakan sebagai dasar untuk ditingkatkan dengan pengintegrasian attention layer. Pendekatan ini bertujuan untuk meningkatkan kemampuan model dalam mengatasi data time series secara lebih akurat dan efisien.

Berikut hasil evaluasi dan prediksi dengan penambahan attention layer pada model LSTM-RNN:



Gambar 4.30 Grafik Confidence Interval MAE pada model LSTM-RNN

A graph of a loss

Description automatically generated

Gambar 4.31 Grafik Confidence Interval *LOSS* pada model LSTM-RNN

Grafik 4.26 dan 4.27 tersebut menampilkan perjalanan dari *history* *loss* dan MAE model RNN-LSTM selama 100 kali iterasi. Garis biru menggambarkan nilai terbesar dari *history* *loss* dan MAE, yang tercapai pada suatu titik selama iterasi. Garis hijau merepresentasikan nilai terkecil dari *history* *loss* dan MAE, yang mencerminkan performa terbaik model atau analisis pada iterasi tertentu. Selanjutnya, rata-rata dari *history* *loss* dan MAE dihitung dan direpresentasikan sebagai garis merah. Dari grafik tersebut juga terlihat bahwa *history* *loss* secara konsisten menurun seiring dengan berjalannya iterasi. Hal ini menunjukkan bahwa model atau analisis yang digunakan secara umum semakin mendekati hasil yang lebih akurat atau optimal seiring berjalannya waktu atau iterasi.

A graph of a graph showing the amount of time

Description automatically generated

Gambar 4.32 Grafik Confidence Interval hasil prediksi pada model LSTM-RNN

Grafik 4.28 tersebut menggambarkan hasil prediksi (garis biru dan garis hijau) dari model LSTM-RNN dengan Attention Layer ketika dijalankan 100 kali. Garis biru mewakili nilai terbesar yang dihasilkan dari prediksi tersebut, sedangkan garis hijau merepresentasikan nilai terkecil dari hasil prediksi. Selanjutnya, hasil prediksi yang diperoleh dari model atau analisis tersebut dicocokkan dengan nilai sebenarnya dan direpresentasikan sebagai garis merah. Berdasarkan visualisasi tersebut, tampak bahwa hasil prediksi (garis biru dan garis hijau) relatif konsisten dan berada di sekitar garis merah (hasil aktual). Hal ini menunjukkan bahwa model atau analisis tersebut memberikan prediksi yang relatif mendekati nilai aktual dari data, dan memiliki tingkat kesalahan yang cukup kecil.

## Implementasi pada Asset Integrity

### Data Petronas

Berikut merupakan data time series dari Petronas:

A graph showing a number of blue lines

Description automatically generated

Gambar 4.32 Data Dissolved Oxygen

Penelitian ini memperlihatkan data tentang kadar oksigen terlarut dalam air yang diukur dalam satuan ppb melalui grafik di atas. Data tersebut mencakup pengamatan harian pada sumbu x, yang menunjukkan perubahan nilai ppb oksigen terlarut pada sumbu y. Satuan ppb digunakan untuk melaporkan jumlah oksigen gas yang larut dalam air. Penelitian ini dapat memberikan informasi penting tentang kualitas air dan potensi keberlanjutan ekosistem perairan yang terkait.

### Preprocessing Data

Langkah pertama dalam preprocessing data time series pada data Petronas adalah dengan melakukan deteksi dan penanganan nilai yang hilang atau data yang hilang. Hal ini dapat dilakukan dengan mengisi nilai yang hilang dengan teknik imputasi seperti interpolasi linear atau menggunakan nilai dari titik data sebelumnya atau sesudahnya. Langkah selanjutnya adalah menangani potensi outliers dalam data. Outliers dapat mempengaruhi analisis secara signifikan, dan oleh karena itu perlu diperiksa dan diberikan penanganan yang sesuai. Salah satu metode yang umum digunakan adalah menggunakan pendekatan statistik, seperti metode deteksi outlier berbasis deviasi standar atau menggunakan metode interkuartil.

### Hasil Evaluasi LSTM-RNN

Berikut hasil evaluasi menggunakan model LSTM-RNN berdasarkan *history* *loss* dan MAE yang dihasilkan selama proses training:

A graph with blue and orange lines

Description automatically generated

Gambar 4.8 Grafik Evaluasi Model LSTM-RNN

Dalam proses analisis evaluasi model LSTM-RNN, terdapat manfaat dalam memisahkan grafik awal pelatihan di mana terjadi penurunan nilai yang signifikan. Hal ini dilakukan untuk memperoleh wawasan yang lebih baik tentang perbandingan proporsi antara MAE dan *loss* serta tingkat kesalahan dan keakuratan model. Oleh karena itu, berikut ini disajikan plot 80% terakhir dari grafik evaluasi tersebut:

A graph with blue and orange lines

Description automatically generated

Gambar 4.9 Grafik Evaluasi Model LSTM-RNN (80%)

Grafik *history* *loss* dan MAE pada model LSTM-RNN tersebut menunjukkan penurunan yang signifikan dalam Mean Absolute Error (MAE) dan *loss* pada awal pelatihan serta kecenderungan penurunan yang berlanjut hingga akhir epoch menandakan adanya kemajuan yang baik dalam proses pembelajaran model. Hasil ini menggambarkan bahwa model mampu menyesuaikan diri dengan data pelatihan dan secara efektif mengurangi kesalahan prediksi. Ketika MAE dan *loss* memiliki selisih yang tidak terlalu besar, hal ini menunjukkan bahwa model tersebut tidak mengalami overfitting secara berlebihan.

### Prediksi dengan Model LSTM-RNN

Berikut adalah hasil prediksi dari model LSTM-RNN menggunakan Mean Absolute Error:

A graph showing the results of a performance

Description automatically generated

Gambar 4.10 Grafik Prediksi Model RNN

MAE yang diperoleh dari hasil prediksi tersebut adalah 4.132318 dalam satuan ppb sesuai dengan data Petronas. tersebut menunjukkan nilai yang tidak terlalu jauh berbeda dengan nilai sebenarnya menandakan bahwa model tersebut memiliki tingkat konsistensi yang baik dalam melakukan prediksi. Namun, terdapat beberapa momen yang menunjukkan nilai error yang cukup siginifkan antara nilai sebenarnya dan hasil prediksi.

A person standing in a tree

Description automatically generated

# BAB V KESIMPULAN DAN SARAN



## Kesimpulan

Berdasarkan Mean Absolute Error (MAE) dari hasil training dan evaluasi, dapat disimpulkan bahwa model hybrid LSTM-RNN menghasilkan performa yang lebih baik daripada model RNN dan LSTM secara individual. Model hybrid tersebut menunjukkan hasil prediksi yang lebih akurat dan konsisten. Hal ini menunjukkan bahwa konsep dan penemuan yang terdapat dalam literature review sesuai dengan praktik yang dilakukan dalam pengembangan model. Literature review yang menggambarkan keunggulan dan keefektifan model hybrid LSTM-RNN dalam memodelkan data sequential ternyata terbukti dalam praktik yang dilakukan.

Dengan demikian, hasil ini memperkuat kesesuaian antara temuan dalam literature review dengan implementasi praktik yang dilakukan. Ini menunjukkan keberhasilan dalam mengaplikasikan pengetahuan teoritis dari literature review ke dalam pengembangan model yang lebih baik dan lebih efektif. Keunggulan model hybrid LSTM-RNN adalah bahwa ia mampu menutupi kelemahan yang mungkin ada pada model RNN dan LSTM secara individual. Model RNN cenderung memiliki masalah vanishing gradient, di mana gradien yang menurun secara eksponensial dapat menyebabkan kesulitan dalam melacak hubungan jangka panjang dalam data sequential. Di sisi lain, model LSTM dirancang khusus untuk mengatasi masalah vanishing gradient, namun dapat memiliki kendala dalam memodelkan hubungan jangka pendek yang sederhana.

Dengan menggunakan pendekatan hybrid LSTM-RNN, model dapat memanfaatkan kekuatan masing-masing arsitektur. Model ini menggabungkan keunggulan RNN dalam melacak hubungan sequential jangka pendek dengan kemampuan LSTM dalam menangani hubungan jangka panjang. Dengan demikian, model hybrid LSTM-RNN mampu mengatasi kendala yang ada pada model RNN dan LSTM secara individual, menghasilkan prediksi yang lebih akurat dan kemampuan adaptasi yang lebih baik terhadap pola data sequential yang kompleks. Dalam praktiknya, penggunaan model hybrid LSTM-RNN membuktikan bahwa pendekatan ini menghasilkan performa yang lebih baik dalam pemodelan data sequential, dengan mengatasi kelemahan yang mungkin dimiliki oleh model RNN atau LSTM secara individual.

Berdasarkan literatur review yang dilakukan terkait penggunaan Attention Layer untuk penelitian ini, tampaknya metode ini tidak berhasil meningkatkan kemampuan prediksi dari model machine learning yang telah dibuat. Meskipun Attention Layer telah digunakan secara luas dalam berbagai aplikasi pemrosesan bahasa alami dan pengenalan pola, ternyata hasilnya tidak sesuai dengan harapan pada penelitian ini. Hal ini menunjukkan bahwa meskipun Attention Layer bisa efektif dalam beberapa konteks, tidak selalu berarti bahwa ia akan memberikan hasil yang baik untuk setiap masalah atau dataset.

## Saran

Meskipun model hybrid LSTM-RNN telah menunjukkan hasil yang lebih baik daripada model standar sebelumnya, ada beberapa saran yang perlu dipertimbangkan untuk penelitian dan pengembangan selanjutnya seperti melakukan pencarian parameter yang lebih cermat untuk memastikan bahwa model diatur dengan optimal. Coba berbagai kombinasi hyperparameter, seperti learning rate, jumlah epoch, dan batch size, serta pertimbangan penggunaan teknik regularisasi seperti dropout atau weight decay. Selain itu, coba uji model pada dataset lain dengan karakteristik time series yang berbeda untuk memastikan bahwa model dapat di generalisasi.

A person standing in a tree

Description automatically generated

# DAFTAR PUSTAKA

Al-Maythalony, B. A., et al. (2019). Predicting corrosion rate of oil and gas pipelines using machine learning algorithms. Journal of Petroleum Science and Engineering, 173, 502-512.

Alghifari, D. R., Edi, M., & Firmansyah, L. (2022). Implementasi Bidirectional LSTM untuk Analisis Sentimen Terhadap Layanan Grab Indonesia. Jurnal Manajemen Informatika (JAMIKA), 12(2), 89-99.

Altan, A., Karasu, S., & Zio, E. (2021). A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. Applied Soft Computing, 100, 106996.

AMAEchi, C. V., Reda, A., Kgosiemang, I. M., Ja’e, I. A., Oyetunji, A. K., Olukolajo, M. A., & Igwe, I. B. (2022). Guidelines on asset management of offshore facilities for monitoring, sustainable maintenance, and safety practices. Sensors, 22(19), 7270.

Asmarini, W. (2023, April 2). Gas Compressor Kilang Pertamina Dumai Riau Meledak, 5 terluka. CNBC Indonesia. https://www.cnbcindonesia.com/news/20230402044850-4-426553/gas-compressor-kilang-pertamina-dumai-riau-meledak-5-terluka

Chen, J., Wang, S., & Wu, X. (2017). Research on carbon reduction of preventive maintenance in manufacturing industry. Procedia Engineering, 174, 548-556.

Dandala, B., Joopudi, V., & Devarakonda, M. (2019). Adverse drug events detection in clinical notes by jointly modeling entities and relations using neural networks. Drug safety, 42, 135-146.

Diskominfotik Riau. (2023, April 2). Media Center. https://mediacenter.riau.go.id/read/77636/kapolda-riau-ungkap-penyebab-ledakan-kilang-p.html

Earthjustice, U. O. C. S. (2022). Joint Labor and Environmental Networks’ Chemical Disaster Prevention Recommendations to the US Chemical Safety and Hazard Investigation Board. NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy, 31(4), 475-483.

Faru, S. H., Waititu, A., & Nderu, L. (2023). A Hybrid Neural Network Model Based on Transfer Learning for Forecasting Forex Market. *Journal of Data Analysis and Information Processing*, *11*(2), 103-120.

Gonzalez, J., & Yu, W. (2018). Non-linear system modeling using LSTM neural networks. IFAC-PapersOnLine, 51(13), 485-489.

Islam, M.S., Hossain, E., Rahman, A., Hossain, M.S. and Andersson, K. (2020) A Review on Recent Advancements in FOREX Currency Prediction. Algorithms, 13, 186. https://doi.org/10.3390/a13080186

Koehrsen, W. (2018). Overfitting vs. underfitting: A complete example. Towards Data Science, 1-12.

Kumar, R., Kumar, P., & Kumar, Y. (2022). Multi-step time series analysis and forecasting strategy using ARIMA and evolutionary algorithms. International Journal of Information Technology, 14(1), 359-373.

Li, J., et al. (2020). Deep learning-based fault diagnosis of rolling bearings considering the timefrequency domain features of vibration signals. Sensors, 20(16), 4586

Liu, F., Cai, M., Wang, L., & Lu, Y. (2019). An ensemble model based on adaptive noise reducer and over-fitting prevention LSTM for multivariate time series forecasting. IEEE Access, 7, 26102-26115

Mahmoud, A., & Mohammed, A. (2021). A survey on deep learning for time-series forecasting. Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges, 365-392.

Mishra, D. S. B. and Alok, D. S. (2011). Handbook Of Research Methodology. Educreation Publishing.

Ningrum, A. A., Syarif, I., Gunawan, A. I., Satriyanto, E., & Muchtar, R. (2021). Algoritma Deep Learning-LSTM untuk Memprediksi Umur Transformator. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 8(3).

Ossai, C. I. (2017). Integrated Big Data Analytics Technique for Real-Time Prognostics, Fault Detection and Identification for Complex Systems. Infrastructures, 2(4), 20.

Pgr. (2023, March 6). Terungkap! Begini Kisah sebelum depo Plumpang Meledak. CNBC Indonesia. https://www.cnbcindonesia.com/news/20230306090832-4-419109/terungkap-begini-kisah-sebelum-depo-plumpang-meledak

Rathore, V.S., Worring, M., Mishra, D.K., Joshi, A. and Maheshwari, S. (2018) Emerging Trends in Expert Applications and Security: Proceedings of ICETEAS 2018. Springer, Berlin, 485-495. https://doi.org/10.1007/978-981-13-2285-3

Raza, J. (2020, August). From Prevent to “Predict & Prevent (PnP)”: Optimizing Oil and Gas Asset Integrity Decisions. In Engineering Assets and Public Infrastructures in the Age of Digitalization: Proceedings of the 13th World Congress on Engineering Asset Management (pp. 511-520). Cham: Springer International Publishing.

Rustam, F., Reshi, A. A., Mehmood, A., Ullah, S., On, B. W., Aslam, W., & Choi, G. S. (2020). COVID-19 future forecasting using supervised machine learning models. IEEE access, 8, 101489-101499.Wang, W., & Lu, Y. (2018, March). Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. In IOP conference series: materials science and engineering (Vol. 324, No. 1, p. 012049). IOP Publishing.

SAKINAH, N., TAHIR, M., BADRIYAH, T., AND SYARIF, I (2019). LSTM with Adam Optimization-Powered High Accuracy Preeclampsia Classification. IES 2019 - International Electronics Symposium: The Role of Techno-Intelligence in Creating an Open Energy System Towards Energy Democracy, Proceedings, 314–319. https://doi.org/10.1109/ELECSYM.2019.8901 536

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. Procedia Computer Science, 181, 526-534.

Sharaf Al-deen, H. S., Zeng, Z., Al-sabri, R., & Hekmat, A. (2021). An improved model for analyzing textual sentiment based on a deep neural network using multi-head attention mechanism. Applied System Innovation, 4(4), 85.

Sun, J., et al. (2019). Pipeline leak detection based on acoustic emission signals using support vector machines. Measurement, 145, 476-485

Sun, J., et al. (2020). Remaining useful life prediction of bearings using an artificial neural network based hybrid approach. Journal of Mechanical Science and Technology, 34(3), 1257-1265

Styp-Rekowski, K., Schmidt, F., & Kao, O. (2020, December). Optimizing Convergence for Iterative Learning of ARIMA for Stationary Time Series. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 2217-2222). IEEE.

Syahrum, S., & Salim, S. (2014). Metodologi Penelitian Kuantitatif.

Tan, H. H., & Lim, K. H. (2019, June). Vanishing gradient mitigation with deep learning neural network optimization. In 2019 7th international conference on smart computing & communications (ICSCC) (pp. 1-4). IEEE.

Tang, K. H. D. (2021). A Case Study of Asset Integrity and Process Safety Management of Major Oil and Gas Companies in Malaysia. Journal of Engineering Research and Reports, 20(2), 6-19.

Tang, Y., et al. (2019). Risk identification and quantitative evaluation method for Asset Integrity management of offshore platform equipment and facilities. Mathematical Problems in Engineering, 2019.

Tian, C., Ma, J., Zhang, C., & Zhan, P. (2018). A deep neural network model for short-term load forecast based on long short-term memory network and convolutional neural network. Energies, 11(12), 3493.

Tian, Y., Zhang, Y., & Zhang, H. (2023). Recent Advances in Stochastic Gradient Descent in Deep Learning. Mathematics, 11(3), 682.

Wang, J., et al. (2018). A comparison study of convolutional neural network and recurrent neural network for bearing remaining useful life prediction. Mechanical Systems and Signal Processing, 107, 494-508.

Xu, L., Li, C., Xie, X., Zhang, G.: Long-short-term memory network based hybrid model for short-term electrical load forecasting. Information 9(7), 165 (2018)

Ying, X. (2019, February). An overview of overfitting and its solutions. In Journal of physics: Conference series (Vol. 1168, p. 022022). IOP Publishing.

Zhou, H., et al. (2021). Remaining useful life prediction of wind turbines based on deep learning models. IEEE Transactions on Sustainable Energy, 12(2), 1032-1042

# LAMPIRAN 1

Parameter Model RNN, LSTM, RNN-LSTM, LSTM-RNN, dan LSTM-RNN

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

# LAMPIRAN 2

A graph with blue and red lines

Description automatically generatedA graph with blue and orange lines

Description automatically generatedEvaluasi Model RNN, LSTM, RNN-LSTM, LSTM-RNN, dan LSTM-RNN

A graph with a line graph

Description automatically generatedA graph with lines and numbers

Description automatically generated

# LAMPIRAN 3

Prediksi Model RNN, LSTM, RNN-LSTM, LSTM-RNN, dan LSTM-RNN

A graph with blue and orange lines

Description automatically generatedA graph with blue and orange lines

Description automatically generated

A graph with blue and orange lines

Description automatically generatedA graph with blue and orange lines

Description automatically generated

# LAMPIRAN 4

Evaluasi (MAE) Model RNN, LSTM, RNN-LSTM, LSTM-RNN, dan LSTM-RNN menggunakan Confidence Interval

A graph of a graph

Description automatically generatedA graph of a graph

Description automatically generatedA graph of a graph

Description automatically generated

A graph of a graph

Description automatically generatedA graph with different colored lines

Description automatically generated

# LAMPIRAN 5

A graph with red and blue lines

Description automatically generatedPrediksi Model RNN, LSTM, RNN-LSTM, LSTM-RNN, LSTM-RNN, LSTM-RNN Attention menggunakan Confidence Interval

A graph of a graph showing the amount of time

Description automatically generatedA graph with red and blue lines

Description automatically generated

A graph with red and blue lines

Description automatically generatedA graph with red and blue lines

Description automatically generated

# LAMPIRAN 6

MAE prediksi Model RNN, LSTM, RNN-LSTM, LSTM-RNN, dan LSTM-RNN menggunakan Confidence Interval