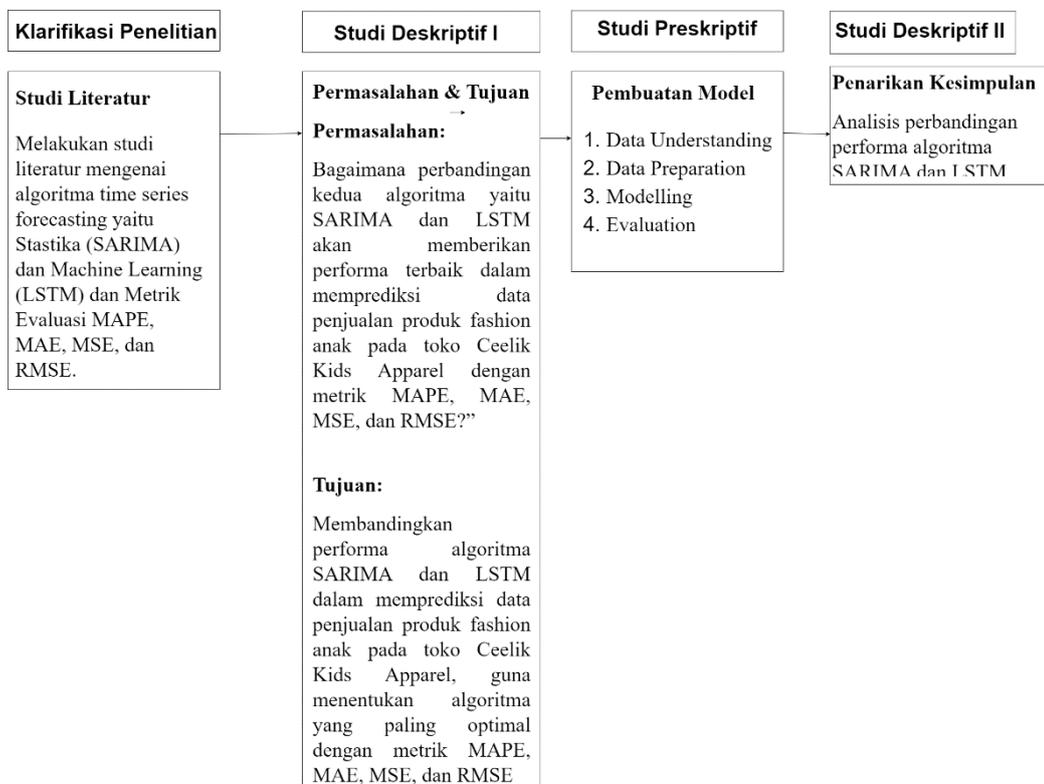


BAB III METODOLOGI PENELITIAN

3.1 Desain Penelitian

Penelitian ini menggunakan pendekatan *Design Research Methodology* (DRM) dalam merancang dan mengembangkan solusi untuk permasalahan yang diidentifikasi. Menurut Blessing-Springer dalam (Putra et al., 2022) DRM menyuguhkan sebuah struktur yang bermanfaat dalam merencanakan penelitian, serta membantu dalam menemukan bidang penelitian secara lebih efisien dan efektif. Pada Gambar 3.1 tahapan DRM dimulai dengan tahap klarifikasi, selanjutnya tahap studi deskriptif I, studi preskriptif dan tahap terakhir, studi deskriptif II.



Gambar 3.1 Desain Penelitian

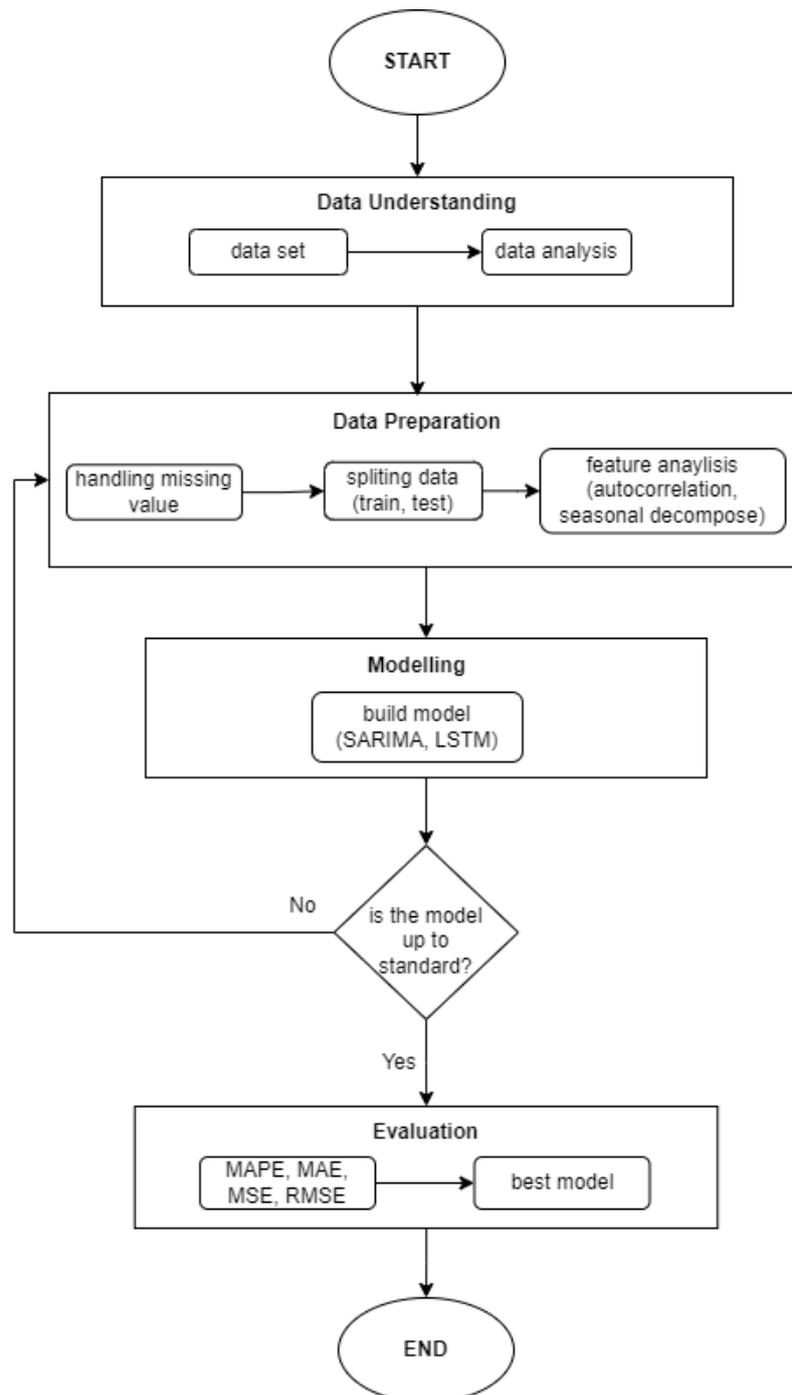
3.1.1 Klarifikasi Penelitian

Tahap pertama dalam penelitian ini melibatkan kegiatan klarifikasi penelitian dengan melakukan studi literatur untuk acuan perencanaan pada penelitian, mengidentifikasi inti masalah, dan menetapkan arah tujuan penelitian yang akan dijalankan. Pengkajian literatur bersumber dari berbagai sumber, seperti jurnal dan buku, yang didasarkan pada penelitian-penelitian sebelumnya melalui akses internet. Literatur yang dikaji mencakup berbagai aspek, mulai dari *e-commerce*, platform *Shopee*, mitra ritel *fashion* yang terlibat dalam penelitian, *data mining*, hingga algoritma yang akan dibandingkan yaitu SARIMA dan LSTM dan metrik evaluasinya yaitu MAPE, MAE, MSE, dan RMSE. SARIMA adalah metode statistik yang diperkenalkan oleh Box dan Jenkins, yang dikenal karena kemampuannya dalam menangani data yang memiliki pola musiman. Menurut Hyndman dan Athanasopoulos (2018), SARIMA adalah salah satu algoritma terbaik untuk menangani deret waktu dengan komponen musiman karena algoritma ini mampu menangkap pola berulang pada interval waktu tertentu. Sedangkan LSTM yang diperkenalkan oleh Hochreiter dan Schmidhuber, merupakan salah satu bentuk jaringan saraf dalam yang dirancang untuk mengatasi masalah *vanishing gradient* yang sering terjadi pada algoritma RNN (Recurrent Neural Networks). LSTM mampu mengatasi keterbatasan algoritma statistik tradisional dengan menangani ketergantungan jangka panjang dalam data sekuensial, menjadikannya sangat efektif untuk tugas prediksi deret waktu yang kompleks (Siami-Namini et al., 2018)

3.1.2 Studi Deskriptif I

Tahap ini untuk merinci permasalahan yang akan diangkat serta menetapkan tujuan penelitian secara lebih mendalam. Pusat perhatian studi deskriptif ini adalah perbandingan algoritma *time series forecasting* dalam konteks prediksi penjualan produk di industri ritel *fashion*. Permasalahan utama yang akan diangkat dalam penelitian ini melibatkan penemuan algoritma *time series forecasting* yang paling efektif untuk mengantisipasi fluktuasi penjualan dalam industri *fashion*. Tujuan penelitian ini berfokus pada identifikasi, perbandingan, dan analisis performa SARIMA dan LSTM.

3.1.3 Studi Preskriptif



Gambar 3.2 Rancangan Model

Gambar 3.2 menunjukkan rancangan model yang dimulai dengan tahap pemahaman data (*Data Understanding*) yaitu dengan mengumpulkan data penjualan produk *fashion* anak-anak dari *Ceelik Kids Apparel* melalui *Shopee Seller Centre*, mencakup periode dari November 2021 hingga November 2023, serta

melakukan analisis data awal. Selanjutnya, tahap persiapan data (*Data Preparation*) mencakup pembersihan data dari *missing values* dan pembagian data menjadi *data training* dan *data testing*. Analisis fitur juga dilakukan, termasuk analisis autokorelasi untuk memahami hubungan antara nilai dalam data *time series* dan dekomposisi musiman (*seasonal decomposition*) untuk memisahkan data menjadi komponen tren, musiman, dan *residual*. Pada tahap pemodelan (*Modelling*), dua algoritma prediksi utama, SARIMA dan LSTM, dibangun dan dilatih menggunakan data yang telah dipersiapkan. Jika model tersebut memenuhi kriteria maka akan dilanjutkan ke tahap evaluasi. Jika tidak memenuhi maka akan dilakukan pengecekan ulang pada tahap persiapan data. Akhirnya, tahap evaluasi (*Evaluation*) dilakukan dengan menggunakan *data testing* untuk mengukur performa algoritma berdasarkan metrik seperti *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE), untuk menentukan algoritma yang terbaik.

3.1.4 Studi Deskriptif II

Tahap terakhir difokuskan pada penarikan kesimpulan. Melakukan analisis perbandingan algoritma SARIMA dan LSTM berdasarkan evaluasi metrik. Analisis mendalam dilakukan untuk mengidentifikasi kekuatan dan kelemahan masing-masing algoritma, dengan menggunakan matrik evaluasi seperti MAPE, MAE, MSE, dan RMSE. Hasil evaluasi algoritma memberikan wawasan tentang kinerja mereka dalam memprediksi *time series* data, dan perbandingan serta pendekatan yang diambil oleh setiap algoritma membantu dalam memahami kondisi di mana masing-masing algoritma menghasilkan performa lebih baik. Rekomendasi yang diberikan didasarkan pada temuan dari perbandingan algoritma, dengan penekanan pada algoritma yang paling sesuai dengan kebutuhan bisnis atau tujuan penelitian.

3.2 Alat Penelitian

Berikut alat dan bahan yang digunakan dalam penelitian ini:

A. Perangkat Keras

1. *Intel(R) Core(TM) i5-5200U CPU @ 2.20GHz 2.20 GHz*
2. *64-bit operating system, x64-based processor*

B. Perangkat Lunak

1. *E-commerce Shopee*
2. *Google Collaboratory*
3. *Python Language Programming*
4. *Python Libraries*

Tabel 3.1
Python Libraries

No	<i>Library</i>	Keterangan
1	<i>numpy</i>	Operasi numerik dan array.
2	<i>pandas</i>	Digunakan untuk manipulasi dan analisis data.
3	<i>matplotlib.pyplot</i>	Digunakan untuk visualisasi data.
4	<i>seaborn</i>	Digunakan untuk visualisasi data yang lebih menarik dan informatif.
5	<i>tensorflow</i>	Digunakan untuk membangun dan melatih algoritma deep learning.
6	<i>os</i>	Digunakan untuk interaksi dengan sistem operasi, seperti manajemen file.
7	<i>datetime</i>	Digunakan untuk manipulasi tanggal dan waktu.
8	<i>sklearn.algoritma_selection.t rain_test_split</i>	Digunakan untuk membagi dataset menjadi data pelatihan dan pengujian.
9	<i>statsalgoritmas.graphics.tsap lots.plot_acf</i>	Digunakan untuk plot fungsi autokorelasi.

10	<i>statsalgoritmas.tsa.seasonal.seasonal_decompose</i>	Digunakan untuk dekomposisi musiman pada data time series.
11	<i>statsalgoritmas.tsa.statespace.sarima.SARIMA</i>	Digunakan untuk membangun algoritma SARIMA.
12	<i>keras.algoritmas.Sequential</i>	Digunakan untuk membuat algoritma sekuensial dalam Keras.
13	<i>keras.layers.LSTM</i>	Digunakan untuk menambahkan lapisan LSTM pada algoritma.
14	<i>keras.layers.Dense</i>	Digunakan untuk menambahkan lapisan dense pada algoritma.
15	<i>sklearn.preprocessing.MinMaxScaler</i>	Digunakan untuk normalisasi fitur.
16	<i>keras.callbacks.EarlyStopping</i>	Digunakan untuk menghentikan pelatihan algoritma lebih awal jika tidak ada perbaikan
17	<i>sklearn.metrics</i>	Menghitung evaluasi

C. Dataset

Dataset yang digunakan dalam penelitian ini diperoleh langsung dari mitra terkait, yaitu data penjualan dari Ceelik Kids Apparel, dengan periode mulai dari November 2021 hingga November 2023. Total data yang dikumpulkan adalah sebanyak 17.144 data, yang tersedia per bulan. Data transaksi ini diambil dari *website* resmi *Shopee Seller Centre*.

3.3 Data Penelitian

Adapun bahan data yang digunakan dalam penelitian ini diperoleh dari mitra toko *Shopee*, yaitu *Ceelik Kids Apparel*. *Dataset* tersebut mencakup data transaksi penjualan produk *fashion* anak dari November 2021 hingga November 2023. Data transaksi ini menjadi komponen kunci dalam pelatihan pada algoritma algoritma SARIMA dan LSTM untuk *forecasting*, memungkinkan penelitian ini untuk menghasilkan prediksi yang akurat. Berikut penjelasan dari masing-masing data yang akan dilakukan analisa yang terdapat pada Tabel 3.2:

Tabel 3.2
Data Penelitian

No	Data yang diambil	Deskripsi	Tujuan
1	Nama Produk	Mengidentifikasi produk <i>fashion</i> anak yang dibeli oleh pembeli setiap transaksi dari <i>Ceelik Kids Apparel</i> melalui platform <i>Shopee</i>	Untuk mengetahui performa penjualan berdasarkan produk.
2	Jumlah	Data ini mencatat jumlah unit produk yang terjual dalam setiap transaksi	Untuk menganalisis volume penjualan produk secara keseluruhan dan melacak perubahan tren penjualan dari waktu ke waktu.
3	Provinsi	Data ini menunjukkan lokasi geografis pembeli berdasarkan provinsi atau wilayah	Untuk mendapatkan insight data dari segmentasi pasar berdasarkan lokasi pembeli
4	Waktu Pembayaran Dilakukan	Data ini mencatat waktu transaksi pembayaran dilakukan oleh pembeli	Untuk menganalisis pola pembelian.

3.4 Instrumen Penelitian

Instrumen penelitian untuk penelitian ini mencakup penggunaan *library* dalam pengembangan dan evaluasi algoritma, serta penerapan rumus-rumus untuk mengukur keberhasilan algoritma. Dalam pengembangan algoritma *time series forecasting* dengan metode kuantitatif seperti SARIMA dan LSTM penelitian ini akan menggunakan bahasa pemrograman *Python* sebagai platform utama untuk pengembangan algoritma.

Evaluasi performa algoritma akan dilakukan dengan memanfaatkan metrik kesalahan seperti *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE). MAPE untuk mengukur kesalahan relatif antara nilai aktual dan nilai prediksi dalam bentuk persentase, MAE mengukur rata-rata dari selisih absolut antara nilai prediksi dan nilai aktual, MSE mengukur rata-rata dari kuadrat selisih antara nilai prediksi dan nilai aktual dan RMSE adalah akar kuadrat dari MSE, memberikan gambaran kesalahan rata-rata dengan satuan yang sama dengan variabel terukur. Metode ini digunakan untuk mengukur tingkat kesalahan algoritma dan kualitas prediksinya, memberikan gambaran yang komprehensif tentang sejauh mana algoritma dapat memprediksi penjualan dengan akurat.

Untuk menganalisis data, langkah-langkah yang akan dilakukan mulai dari pengumpulan data penjualan dari *Ceelik Kids Apparel* akan dilakukan dari *Shopee Seller Centre*, dengan periode mulai dari November 2021 hingga November 2023. Data ini akan dibagi menjadi data pelatihan (data *training*) dan data pengujian (data *testing*) untuk membangun dan menguji algoritma. Selanjutnya, data akan diproses dengan membersihkan *missing values*. Analisis eksploratif akan dilakukan untuk memahami pola dan tren dalam data menggunakan visualisasi seperti plot *time series*. Algoritma *time series forecasting* kemudian akan dikembangkan menggunakan algoritma SARIMA dan LSTM dengan *library Python*. Evaluasi algoritma dilakukan menggunakan data pengujian dan metrik kesalahan yaitu MAPE, MAE, MSE, dan RMSE. Hasil evaluasi akan dianalisis untuk memahami seberapa baik algoritma memprediksi penjualan dan faktor-faktor yang mempengaruhi performa algoritma. Terakhir, kesimpulan dan rekomendasi akan disusun berdasarkan hasil analisis dan evaluasi algoritma.