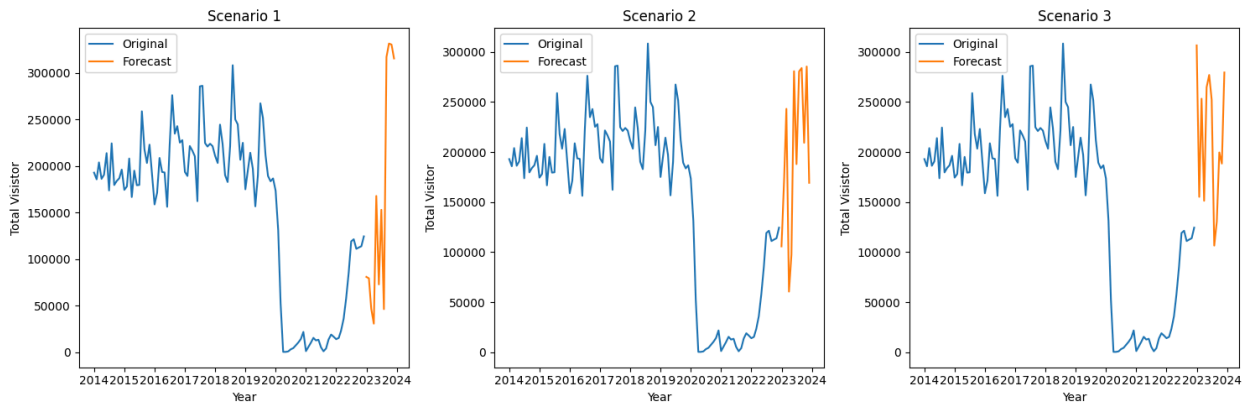


# Hasil Forecast

Berikut merupakan hasil grafik visualisasi 12 periode kedepan untuk total\_visitor dengan menggunakan variabel total\_accomodation sebagai faktor yang mempengaruhi. Dibawah disajikan pula ringkasan metrik evaluasi dari masing-masing scenario.

## 1. Recurrent Neural Network (RNN)

Hasil Grafik Forecast dengan RNN:



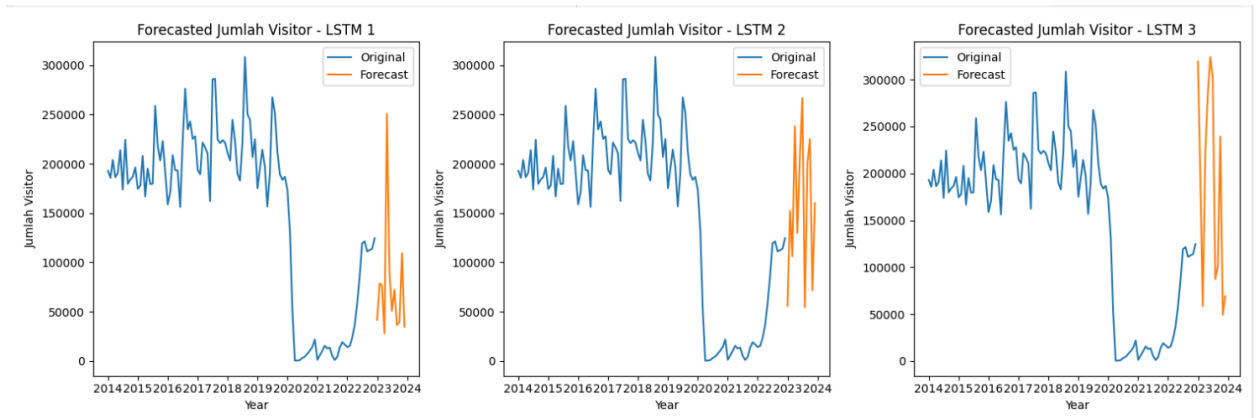
Hasil ringkasan metrik evaluasi:

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	RNN	0.0112279	0.12148	0.0738795	0.105962	0.883677
Scenario 2	RNN	0.00372106	0.100705	0.0516899	0.0610005	0.932781
Scenario 3	RNN	0.0384279	0.4931	0.186455	0.19603	0.260712

Dari hasil ringkasan metrik evaluasi di atas dapat disimpulkan bahwa **Scenario 2** dengan pembagian dataset 70% training dan 30% testing memiliki performa yang lebih baik diantara skenario yang lain. Hal itu dapat terlihat dari nilai MSE, MAPE, MAE, dan RMSE yang lebih kecil daripada skenario yang lain, dan juga memiliki nilai R kuadrat yang paling tinggi.

## 2. Long Short Term Memory (LSTM)

Hasil Grafik Forecast dengan LSTM



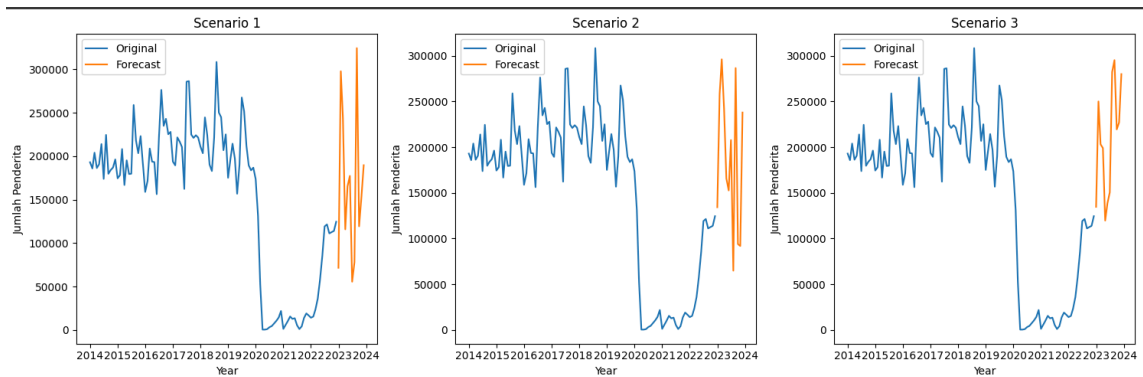
Hasil ringkasan metrik evaluasi:

Scenario	Model	RMSE	MSE	MAPE	MAE	R2
Scenario 1	LSTM	0.0489583	0.00239692	0.149016	0.0413247	0.94253
Scenario 2	LSTM	0.0455629	0.00207598	0.15067	0.039771	0.971855
Scenario 3	LSTM	0.0455629	0.00432381	0.226786	0.0640599	0.960643

Terlihat bahwa **Scenario 2** dengan 70% data train dan 30% data test memiliki performa yang lebih baik. Hal ini dibuktikan dari nilai MSE dan MAE yang lebih kecil dibandingkan scenario lainnya. Selain itu, nilai r-squared nya juga yang tertinggi atau paling mendekati 1 yang mengindikasikan bahwa model yang dilatih tergolong bagus. Sehingga dapat disimpulkan untuk Scenario terbaik dalam menggunakan model LSTM pada dataset Jakarta adalah dengan 70% Training dan 30% Testing.

### 3. Gated Recurrent Unit (GRU)

Hasil Grafik Forecast dengan GRU pada ketiga Scenario



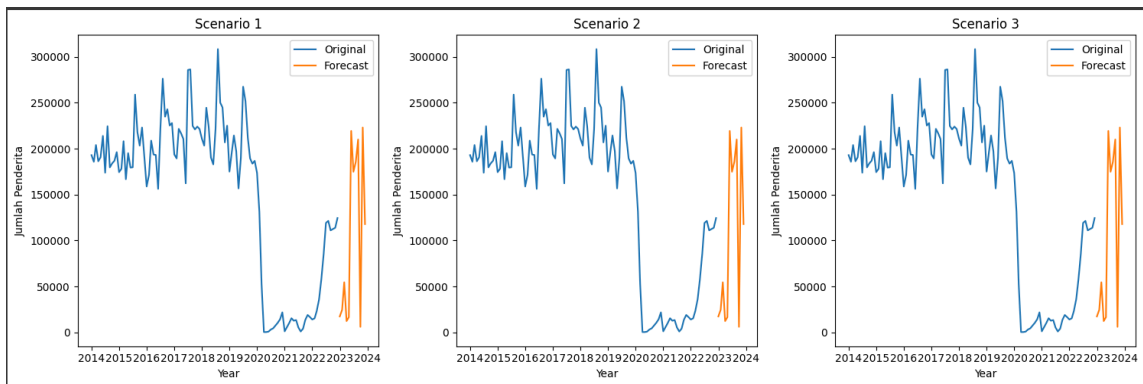
Hasil Evaluasi Model GRU pada ketiga Scenario

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	GRU	0.00353303	0.0856239	0.0548102	0.0594393	0.926345
Scenario 2	GRU	0.00306259	0.187444	0.0426237	0.0553407	0.963181
Scenario 3	GRU	0.00213531	0.0836964	0.0392766	0.0462094	0.955484

Dari hasil tersebut, terlihat bahwa **Scenario 3** memiliki MAPE, MSE, RMSE yang paling rendah, yang menunjukkan bahwa akurasi prediksinya lebih baik dibandingkan scenario lainnya. Pada Scenario 3 juga memiliki nilai  $R^2$  yang mendekati 1 yang menunjukkan model cocok dengan data. Sehingga dapat disimpulkan untuk Scenario terbaik dalam menggunakan model GRU pada dataset Jakarta adalah dengan 60% Training dan 40% Testing.

#### 4. Support Vector Regression (SVR)

Hasil Grafik Forecast dengan SVR pada ketiga scenario



Hasil Evaluasi Model SVR pada ketiga Scenario

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661
Scenario 2	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661
Scenario 3	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661

Dari tabel hasil evaluasi, terlihat bahwa nilai MSE, MAPE, MAE, dan R2 pada ketiga scenario memiliki nilai yang tidak memiliki perbedaan yang signifikan. Hal ini dikarenakan data yang digunakan memiliki pola yang sama. Kami tidak menggunakan shuffled data dikarenakan data berbentuk time series, sehingga dengan pola yang ada model SVR akan mempelajari pola dengan baik. Maka dari itu nilai dari MSE, MAPE dan MAE rendah sehingga nilai target cukup akurat. Faktor lainnya adalah karena penggunaan Kernel linear yang tidak sensitif terhadap skala data. Kernel linear dapat

menghasilkan forecast yang akurat meskipun data training dan testing memiliki skala yang berbeda.

## I. Hasil Analisis dan Kesimpulan

Berdasarkan 4 metode forecasting yang telah digunakan dengan percobaan pada 3 skenario pembagian dataset latihan dan uji, dapat diambil kesimpulan sebagai berikut:

- **Metode RNN** terbaik menggunakan skenario 2 dengan proporsi pembagian dataset 70% training dan 30% testing karena nilai MSE, MAPE, MAE, dan RMSE yang lebih kecil daripada skenario yang lain, dan juga memiliki nilai R kuadrat yang paling tinggi. Model RNN ini menggunakan data latihan 12 bulan sebelumnya untuk prediksi 1 data setelahnya dan menggunakan 3 layer RNN dengan fungsi aktivasi tanh untuk 1 output layer
- **Metode LSTM** terbaik menggunakan skenario 2 dengan proporsi pembagian dataset 70% training dan 30% testing karena nilai MSE dan MAE yang lebih kecil daripada skenario yang lain, dan juga memiliki nilai R kuadrat yang paling tinggi. Model LSTM ini juga menggunakan data latihan 12 bulan sebelumnya untuk prediksi 1 data setelahnya dan menggunakan 3 layer LSTM dengan fungsi aktivasi tanh untuk 1 output layer
- **Metode GRU** terbaik menggunakan skenario 3 dengan proporsi pembagian dataset 60% training dan 40% testing karena nilai MSE, RMSE dan MAPE yang lebih kecil daripada skenario yang lain, dan juga memiliki nilai R kuadrat yang tinggi mendekati nilai 1. Model GRU ini menggunakan data latihan 12 bulan sebelumnya untuk prediksi 1 data setelahnya dan menggunakan 3 layer GRU dengan fungsi aktivasi 'tanh' untuk 1 output layer
- **Metode SVR** menghasilkan hasil evaluasi yang sama diantara 3 skenario karena metode SVR tidak sensitif terhadap skala data sehingga tidak dipengaruhi oleh pembagian data latihan dan data uji.

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	RNN	0.0112279	0.12148	0.0738795	0.105962	0.883677
Scenario 2	RNN	0.00372106	0.100705	0.0516899	0.0610005	0.932781
Scenario 3	RNN	0.0384279	0.4931	0.186455	0.19603	0.260712

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	GRU	0.00353303	0.0856239	0.0548102	0.0594393	0.926345
Scenario 2	GRU	0.00306259	0.187444	0.0426237	0.0553407	0.963181
Scenario 3	GRU	0.00213531	0.0836964	0.0392766	0.0462094	0.955484

Scenario	Model	RMSE	MSE	MAPE	MAE	R2
Scenario 1	LSTM	0.0489583	0.00239692	0.149016	0.0413247	0.94253
Scenario 2	LSTM	0.0455629	0.00207598	0.15067	0.039771	0.971855
Scenario 3	LSTM	0.0455629	0.00432381	0.226786	0.0640599	0.960643

Scenario	Model	MSE	MAPE	MAE	RMSE	R2
Scenario 1	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661
Scenario 2	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661
Scenario 3	SVR	2.80469e-05	0.0987418	0.00493044	0.00529593	0.999661

- Dari ketiga metode deep learning antara RNN, GRU dan LSTM, dilihat dari nilai MSE terkecil, didapatkan bahwa **model LSTM** memiliki nilai MSE paling rendah.
- LSTM dan GRU juga merupakan model yang lebih efektif daripada RNN dalam menangani data urutan kompleks dan panjang karena kemampuan mereka untuk memahami dan memanipulasi informasi jangka panjang yang cocok untuk dataset yang diberikan dalam jangka waktu tahun 2014 hingga 2022
- Untuk model terbaik dari keempat metode yang kami buat adalah pada **metode SVR**, ditunjukkan dari nilai MSE yang paling kecil diantara keempat metode. Hal ini dikarenakan pada dataset Jakarta dengan data total\_visitor dan total\_accomodation merupakan jenis data nonlinear seperti yang terlihat pada grafik di bawah sehingga dengan menggunakan kernel linear akan mentransformasikan ke ruang dimensi yang lebih tinggi dan SVR akan mampu membaca pola data dengan lebih baik.

```

# Menggambarkan plot untuk total_visitor (warna biru)
plt.plot(data_asli.index, data_asli['total_visitor'], color='blue', label='Total Visitor')

# Menggambarkan plot untuk total_accommodation (warna merah)
plt.plot(data_asli.index, data_asli['total_accommodation'], color='red', label='Total Accommodation')

# Menambahkan label pada sumbu x dan y
plt.xlabel('Tanggal')
plt.ylabel('Jumlah')

# Menambahkan judul plot
plt.title('Total Visitor dan Total Accommodation Seiring Waktu - BEFORE NORMALIZATION')

# Menambahkan legenda
plt.legend()

# Menampilkan grid lines
plt.grid(True)

# Menampilkan plot
plt.show()

```

Total Visitor dan Total Accommodation Seiring Waktu - BEFORE NORMALIZATION

