



# Analisis Peramalan dengan *Long Short Term Memory* pada Data Kasus Covid-19 di Provinsi Jawa Tengah

Putri Anis Qori<sup>a,\*</sup>, Dhama Sekar Oktafani<sup>b</sup>, Iqbal Kharisudin<sup>a,b</sup>

<sup>a</sup> Universitas Negeri Semarang, Semarang 50229, Indonesia

\* Alamat Surel: [putria1@students.unnes.ac.id](mailto:putria1@students.unnes.ac.id)

## Abstrak

Makalah ini bertujuan untuk menjelaskan model *Long Short Term Memory* pada penerapan analisis peramalan kasus Covid-19 di Provinsi Jawa Tengah. Penelitian ini menggunakan google colab dengan bahasa pemrograman python. Data yang diamati adalah data harian dari 8 Maret 2020 sampai dengan 6 September 2021 sebanyak 548 periode. Hasil analisis time series menunjukkan bahwa model terbaik untuk meramal kasus Covid-19 di Provinsi Jawa Tengah adalah metode LSTM untuk regresi menggunakan metode window. Hal ini dikarenakan ketepatan nilai RMSE terendah sebesar 715,62 dibandingkan dengan metode yang lainnya. Dalam peramalan ini terdapat kemungkinan bahwa sewaktu-waktu berubah dikarenakan faktor tertentu. Kasus pandemi ini juga bisa berubah sewaktu-waktu tergantung upaya pemerintah Provinsi Jawa Tengah dan juga kesadaran masyarakat untuk selalu menjaga kebersihan dan menerapkan protokol kesehatan.

## Kata kunci:

Peramalan, Covid-19, *long short term memory*, Python

© 2022 Dipublikasikan oleh Jurusan Matematika, Universitas Negeri Semarang

## 1. Pendahuluan

### 1.1. Latar Belakang

Indonesia dilanda wabah virus corona pada awal Maret lalu. Prevalensi Covid-19 di Indonesia terus meningkat dari hari ke hari. Pemerintah menghimbau kepada masyarakat untuk melakukan social distance untuk memutus mata rantai penyebaran wabah Covid-19 yang semakin meningkat. Saat terjadi kasus corona, penulis membutuhkan data yang akurat untuk dapat menanganinya secara cepat dan tepat. Selain itu, kita harus mampu mengantisipasi apa yang akan terjadi di masa depan dan menunjukkan apa yang akan terjadi. Prakiraan ini harus menunjukkan bahwa hasil prakiraan realistis sehingga tidak berbeda secara signifikan dengan tren data aslinya.

Berbagai penelitian masih terus dilakukan dengan tujuan untuk memprediksi kejadian di masa depan dengan akurasi tinggi. Metode statistik tradisional masih banyak digunakan di era kecerdasan buatan (Artificial Intelligence) saat ini metode machine learning (machine learning) telah banyak berkembang dalam berbagai bidang. Juga dalam hal peramalan data yang bersifat time series (Karno, 2020). Machine learning adalah algoritma yang meniru cara kerja otak manusia. Mesin ini bisa belajar dari peristiwa masa lalu. Otak manusia menerima banyak informasi secara terus menerus dan berulang-ulang, sehingga semakin banyak data yang kita pelajari, semakin cerdas mesin tersebut. Orang dapat berpikir lebih akurat untuk memecahkan berbagai masalah.

Salah satu metode yang menerapkan machine learning adalah neural network. Neural network adalah model yang dibentuk teinspirasi dari ide cara kerja otak manusia. Salah satu model di neural network untuk menganalisis masalah peramalan adalah Recurrent Neural Network, namun seiring dengan pengembangan pemodelan dengan menggunakan Recurrent Neural Network, pada tahun 1997 mulai dikembangkan model baru yang merupakan versi perubahan perbaikan dari RNN untuk menangani masalah ketergantungan jangka panjang, model ini dikenal dengan Long Short-Term memory (LSTM).

To cite this article:

Qori, P. A., Oktafani, D. S., & Kharisudin, I. (20192). Analisis Peramalan dengan *Long Short Term Memory* pada Data Kasus Covid-19 di Provinsi Jawa Tengah. *PRISMA, Prosiding Seminar Nasional Matematika* 5, 752-758

### 1.2. Rumusan Masalah

Pertumbuhan kasus Covid-19 selalu berkembang seiring waktu, oleh karena itu perlu dilakukan peramalan untuk mengetahui kondisi dan mempersiapkan diri dalam menghadapi kasus Covid-19 yang terjadi di periode selanjutnya. Melalui data kasus Covid-19 di Provinsi Jawa Tengah, bagaimana analisis *long short term memory* untuk meramal data kasus Covid-19 di Provinsi Jawa Tengah dan bagaimana hasil keakuratannya.

### 1.3. Tujuan

Dari makalah ini diharapkan dapat melakukan peramalan data kasus Covid-19 di Provinsi Jawa Tengah periode mendatang dan memilih metode terbaik dengan membandingkan hasil keakuratan model.

---

## 2. Metode (untuk artikel hasil kajian, bagian ini tidak ada)

### 2.1. Covid-19

Coronavirus (Covid-19) adalah penyakit menular yang disebabkan oleh sindrom pernapasan akut parah coronavirus 2 (SARSCoV-2) (World Health Organization, 2020). Tanda dan gejala umum infeksi Covid-19 antara lain gejala sesak napas akut seperti demam, batuk, dan sesak napas. Masa inkubasi rata-rata adalah 56 hari dan masa inkubasi terlama adalah 14 hari. Pada kasus Covid-19 yang parah, pneumonia, sindrom pernafasan akut, gagal ginjal bahkan kematian dapat terjadi ("Coronavirus," n.d.) Pada awal Maret 2020, Indonesia dilanda pandemi Covid-19. Pemerintah dan masyarakat menutup jarak sosial untuk memutus rantai Covid-19 yang terus meningkat. Infeksi Covid-19 hanya bisa dicegah dengan disiplin di rumah saja (Sindi et al., 2020).

### 2.2. Peramalan

Peramalan merupakan kegiatan meramal nilai-nilai sebuah peubah kepada nilai yang diketahui dari peubah tersebut atau peubah yang berhubungan (Makridakis, 1999). Sedangkan menurut Subagyo pada tahun 1986, memperkirakan apa yang akan terjadi di masa depan, dan perencanaan menentukan apa yang akan terjadi di masa depan. Peramalan bertujuan untuk mendapatkan peramalan yang dapat meminimalkan kesalahan peramalan.

Berdasarkan penelitian terdahulu penulis dapat mengambil kesimpulan bahwa pengertian peramalan adalah memperamalan sesuatu yang belum terjadi dengan tujuan untuk memperkirakan peristiwa-peristiwa yang akan terjadi dimasa depan dengan memperhatikan data-data dari masa lalu. Peramalan juga dapat didasarkan pada keahlian penilaian, yang pada gilirannya didasarkan pada data historis dan pengalaman (Yuniastari & Wirawan, 2014).

### 2.3. Data Time Series

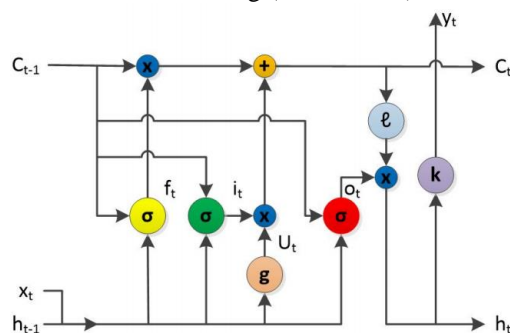
Data time series merupakan data yang dikumpulkan dari waktu ke waktu untuk memberikan gambaran perkembangan kegiatan dari waktu ke waktu. Analisis deret waktu memungkinkan Anda melihat bagaimana satu atau lebih peristiwa terjadi dan bagaimana mereka berhubungan dengan peristiwa lain (Rahmadayanti, Susilo, & Puspitaningrum, 2015). Data time series ada di mana-mana, seperti pada grafik saham, ramalan cuaca, atau pertumbuhan populasi penduduk Metode time series Adalah prediksi kuantitatif berdasarkan analisis pola hubungan antara variabel yang akan dicari (dependensi) dan variabel yang mempengaruhinya (independen), dan perubahan dari waktu ke waktu seperti minggu, bulan, kuartal, kuartal, semester, tahun, dll Untuk melakukan.

Tujuan dari metode ini adalah untuk menemukan pola-pola pada deret masa lalu dan mengekstrapolasi pola-pola tersebut di masa yang akan datang sehingga hasilnya dapat digunakan sebagai acuan untuk memprediksi nilai-nilai yang akan datang..

### 2.4. Long Short Term Memory

*Long Short Term Memory* merupakan sebuah evolusi dari arsitektur RNN, dimana pertama kali diperkenalkan oleh Hochreiter & Schmidhuber pada tahun 1997. Jenis jaringan saraf yang kuat yang dirancang untuk menangani ketergantungan urutan disebut jaringan saraf berulang. Jaringan Long Short-Term Memory atau jaringan LSTM adalah jenis jaringan saraf berulang yang digunakan dalam

pembelajaran mendalam karena arsitektur yang sangat besar dapat berhasil dilatih (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.). Hingga penelitian ini dilakukan banyak para peneliti yang terus mengembangkan arsitektur LSTM di berbagai bidang seperti dalam bidang speech, recognition, dan forecasting. LSTM adalah salah satu jenis dari RNN. LSTM akan menyimpan informasi dari pola pada data. LSTM memiliki kelebihan jika dibandingkan dengan RNN biasa. LSTM dapat mempelajari sebuah data yang mana nantinya jika hasil perhitungan data adalah 1 maka akan disimpan dan apabila nilai data 0 maka akan dibuang (Johan, 2019).



**Gambar 1.** Desain *Long Short Term Memory*

Neuron pada LSTM terdiri dari forget gates, input gates, new cell state, dan output gates yang dapat mengatur memori pada masing-masing neuron itu sendiri sehingga dapat memproses data time series yang jumlahnya sangat banyak (Zheng, Xu, Zhang, & Li, 2017). Setiap keluaran mempunyai fungsinya masing-masing; forget gates: secara kondisional memutuskan informasi apa yang akan dibuang dari blok, sedangkan input gates: secara kondisional memutuskan nilai mana dari input untuk memperbarui status memori, serta output gates: secara kondisional memutuskan apa yang akan dikeluarkan berdasarkan masukan dan memori blok. Setiap unit seperti mesin mini-state di mana gerbang unit memiliki bobot yang dipelajari selama prosedur pelatihan.

#### 2.4.1. LSTM untuk Regresi

Pada analisis LSTM untuk regresi kita dapat merumuskan masalah sebagai masalah regresi. Menulis fungsi sederhana untuk mengubah satu kolom data menjadi kumpulan data dua kolom: kolom pertama berisi ( $t$ ) jumlah penumpang bulan ini dan kolom kedua berisi ( $t+1$ ) jumlah penumpang bulan depan, yang akan diprediksi (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.).

#### 2.4.2. LSTM untuk Regresi Menggunakan Metode Window

Pada analisis LSTM untuk regresi menggunakan metode window kami juga dapat merumuskan masalah sehingga beberapa langkah waktu terbaru dapat digunakan untuk membuat prediksi untuk langkah waktu berikutnya yang disebut jendela, dan ukuran jendela adalah parameter yang dapat disesuaikan untuk seper masalah. Misalnya, mengingat waktu sekarang ( $t$ ) kita ingin memprediksi nilai pada waktu berikutnya dalam urutan ( $t+1$ ), kita dapat menggunakan waktu sekarang ( $t$ ), serta dua waktu sebelumnya ( $t-1$  dan  $t-2$ ) sebagai variabel input. Ketika diutarakan sebagai masalah regresi, variabel inputnya adalah  $t-2$ ,  $t-1$ ,  $t$  dan variabel outputnya adalah  $t+1$  (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.).

#### 2.4.3. LSTM untuk Regresi dengan Langkah Waktu

Pada analisis LSTM untuk regresi dengan langkah waktu bahwa persiapan data untuk jaringan LSTM mencakup langkah-langkah waktu. Beberapa masalah urutan mungkin memiliki jumlah langkah waktu yang bervariasi per sampel. Misalnya, Anda mungkin memiliki pengukuran mesin fisik yang mengarah ke titik kegagalan atau titik lonjakan. Seper kejadian akan menjadi sampel pengamatan yang mengarah ke peristiwa tersebut akan menjadi langkah waktu, dan variabel yang diamati akan menjadi fitur. Langkah waktu menyediakan cara lain untuk mengungkapkan masalah deret waktu kita. Seperti contoh jendela di atas, kita dapat mengambil langkah waktu sebelumnya dalam deret waktu kita sebagai input untuk memprediksi output pada langkah waktu berikutnya. Alih-alih mengungkapkan pengamatan masa lalu sebagai fitur input terpisah, kita dapat menggunakannya sebagai langkah waktu dari satu fitur input, yang

memang merupakan pembedaan masalah yang lebih akurat. Kita dapat melakukan ini dengan menggunakan representasi data yang sama seperti pada contoh berbasis jendela sebelumnya (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.).

#### 2.4.4. LSTM untuk Regresi dengan Langkah Waktu

Pada analisis LSTM memiliki memori, yang mampu mengingat seluruh urutan yang panjang. Biasanya, status dalam jaringan disetel ulang setelah seper batch pelatihan saat memasang model, serta seper panggilan ke model.predict() atau model.evaluate(). Kita bisa mendapatkan kontrol yang lebih baik ketika keadaan internal jaringan LSTM dibersihkan di Keras dengan membuat lapisan LSTM "status". Ini berarti bahwa ia dapat membangun status di seluruh urutan pelatihan dan bahkan mempertahankan status itu jika diperlukan untuk membuat prediksi. Ini mengharuskan data pelatihan tidak dikocok saat memasang jaringan. Ini juga memerlukan pengaturan ulang eksplisit dari status jaringan setelah seper pemaparan ke data pelatihan (Epoch) dengan panggilan ke model.reset\_states(). Ini berarti bahwa kita harus membuat loop luar epoch kita sendiri dan di dalam seper epoch panggil model.fit() dan model.reset\_states() (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.).

#### 2.4.5. LSTM untuk Regresi dengan Memori Antar Batch

Terakhir pada analisis LSTM ditumpuk dengan memori antar batch kita akan melihat salah satu manfaat besar LSTM: fakta bahwa mereka dapat berhasil dilatih saat ditumpuk ke dalam arsitektur jaringan dalam. Jaringan LSTM dapat ditumpuk di Keras dengan cara yang sama seperti jenis lapisan lainnya dapat ditumpuk. Satu tambahan untuk konfigurasi yang diperlukan adalah bahwa lapisan LSTM sebelum seper lapisan LSTM berikutnya harus mengembalikan urutannya. Ini dapat dilakukan dengan menyetel parameter return\_sequences pada layer ke True. Kita dapat memperluas LSTM stateful di bagian sebelumnya agar memiliki dua lapisan (“Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras,” n.d.).

### 2.5. Menghitung Nilai Akurasi

RMSE merupakan alternatif dari proses peramalan. Pendekatan ini penting karena pendekatan ini cukup rawan kesalahan dan mengutamakan prediksi yang menghasilkan kesalahan yang sangat besar. Metode ini sering digunakan untuk mengevaluasi hasil ramalan. Saat menggunakan MSE, kesalahan menunjukkan seberapa besar perbedaan antara hasil estimasi dan hasil estimasi. Bedanya, datanya acak atau tidak mengandung perkiraan lagi. Nilai RMSE dapat dihitung dengan rumus (Yuniastari & Wirawan, 2014) sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_i^n (\hat{x}_i - x_i)^2}{n}} \quad (1)$$

dengan,

$\hat{x}$  = nilai hasil peramalan

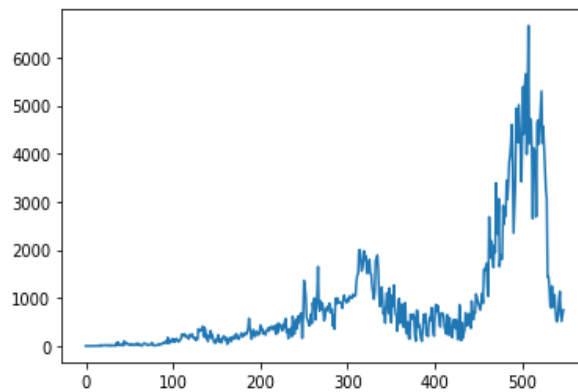
$x_i$  = nilai observasi ke-i

$n$  = banyaknya data

---

## 3. Hasil dan Pembahasan

### 3.1. Grafik Tren Covid-19 di Provinsi Jawa Tengah



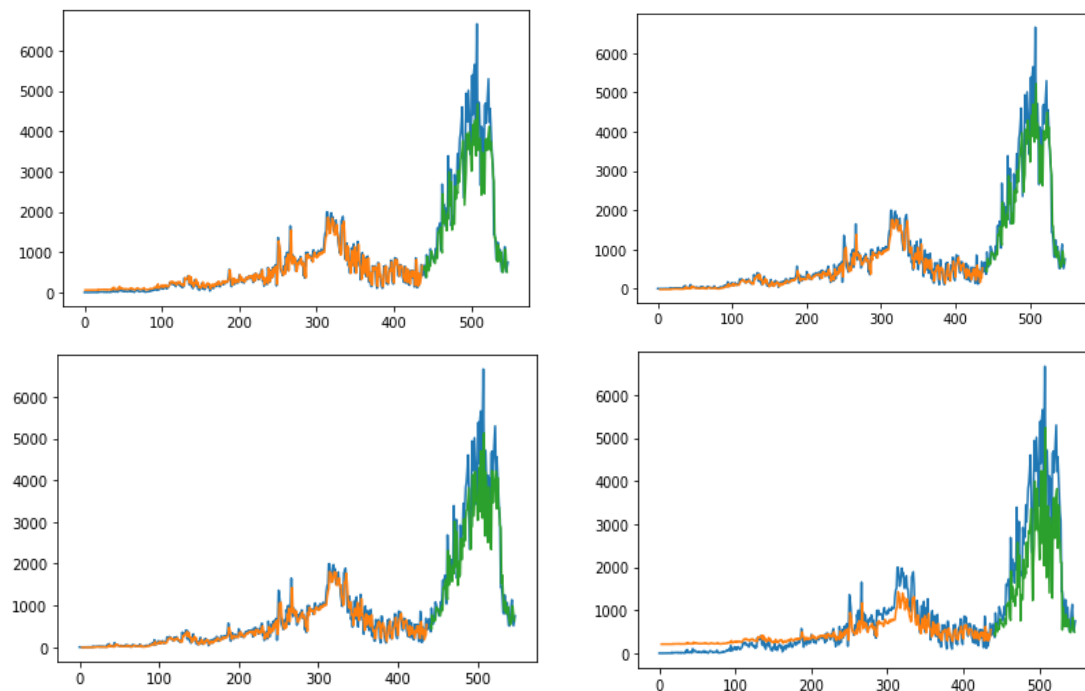
**Gambar 2.** Pergerakan Kasus Covid-19 di Provinsi Jawa Tengah

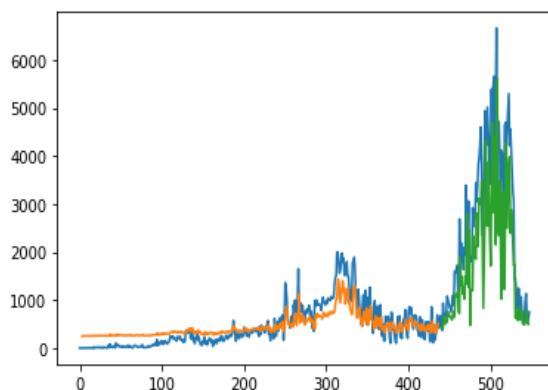
Berdasarkan gambar diatas diperoleh informasi bahwa pergerakan kasus Covid-19 di Provinsi Jawa Tengah memiliki pola trend kenaikan dan sangat fluktuatif.

### 3.2. Normalisasi Data dan Splitting Data

Data yang dihasilkan oleh analisis peramalan memiliki nilai yang bervariasi, oleh sebab itu untuk meminimalisir error maka dilakukan normalisasi data. Pada penelitian ini metode yang digunakan dalam uji normalitas adalah metode *Min-Max Scaler*.

Selanjutnya dibutuhkan data latih dan data uji, pembagian data ini digunakan untuk membedakan antara data yang digunakan untuk proses pelatihan dan data yang digunakan untuk proses pengujian. Data pelatihan dibuat untuk meningkatkan kinerja LSTM dalam pengujian data, sehingga memilih parameter terbaik untuk model. Seperti yang ditunjukkan pada Tabel 1, total 80 juta total data akan digunakan sebagai data latih dan 20.000 total data akan digunakan sebagai data uji. Jumlah data latih akan lebih banyak, dengan tujuan untuk melatih lebih baik mesin latih untuk melatih model. Ini memberikan prediksi data uji yang lebih baik saat mesin menghasilkan model. Model LSTM Kemudian dibuatlah model untuk masing masing metode. Setelah model dengan data training terbentuk maka dilakukan pengujian model terhadap data testing.





**Gambar 4.** (a) LSTM untuk regresi; (b) LSTM untuk regresi menggunakan metode window; (c) LSTM untuk regresi dengan langkah waktu; (d) LSTM dengan memori antar batch; (e) LSTM ditumpuk dengan memori antar batch

### 3.3. Memilih Model Terbaik

Dalam proses ini didapatkan RMSE (Root Mean Square Error) atau nilai keakuratan masing-masing metode.

**Tabel 1.** Nilai keakuratan masing-masing metode

Metode	RMSE
LSTM untuk regresi	787,87
LSTM untuk regresi menggunakan metode window	715,62
LSTM untuk regresi dengan langkah waktu	796,91
LSTM dengan memori antar batch	1156,65
LSTM ditumpuk dengan memori antar batch	1218,12

Berdasarkan Tabel 1 didapatkan nilai RMSE terendah adalah 715,62 dengan menggunakan metode LSTM untuk regresi menggunakan metode window.

## 4. Simpulan

Pada hasil peramalan didapatkan kesimpulan bahwa data kasus Covid-19 di Provinsi Jawa Tengah memiliki tren naik. Berdasarkan hasil pembahasan maka diperoleh metode terbaik untuk meramalan data kasus Covid-19 di Provinsi Jawa Tengah adalah metode LSTM untuk regresi menggunakan metode window. Hal ini dikarenakan ketepatan nilai RMSE terendah sebesar 715,62 dibandingkan dengan metode yang lainnya. Dalam peramalan ini terdapat kemungkinan bahwa sewaktu-waktu berubah dikarenakan faktor tertentu.

## Daftar Pustaka

- Johan, K. (2019). *Rancang Bangun Aplikasi*. 7–19.
- Karno, A. S. B. (2020). Prediksi Data Time Series Saham Bank BRI Dengan Mesin Belajar LSTM (Long Short Term Memory). *Journal of Information and Information Security (JIFORTY)*, 1(1), 1–8.
- Makridakis, S. G. (1999). *Forecasting, Planning, and Strategy for the 21 st Century*.
- Rahmadayanti, R., Susilo, B., & Puspitaningrum, D. (2015). *PERBANDINGAN KEAKURATAN METODE AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) DAN EXPONENTIAL SMOOTHING PADA PERAMALAN PENJUALAN SEMEN DI PT. SINAR ABADI*. 3(1), 23–36.
- Sindi, S., Ratnasari, W., Ningse, O., Sihombing, I. A., Zer, F. I. R. H., Hartama, D., & Kunci, K. (2020). *Analisis algoritma k-medoids clustering dalam pengelompokan penyebaran covid-19 di indonesia*.

4(1), 166–173.

Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras. (n.d.). Retrieved September 14, 2021, from <https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>

World Health Organization. (2020). Situation Report 10: Coronavirus Disease 2019 (Covid-19) in Indonesia. *World Health Organization, 19*(June), 1–20.

Yuniastari, N. L., & Wirawan, I. (2014). Peramalan Permintaan Produk Perak Menggunakan Metode Simple Moving Average Dan Exponential Smoothing. *JURNAL SISTEM DAN INFORMATIKA*, 9(1), 97–106.

Zheng, J., Xu, C., Zhang, Z., & Li, X. (2017). *Electric Load Forecasting in Smart Grids Using Long-Short-Term-Memory based Recurrent Neural Network*. 1–6.