

Prediksi Harga Minyak Goreng Curah dan Kemasan Menggunakan Algoritme *Long Short-Term Memory* (LSTM)

Prediction of Bulk and Packaged Cooking Oil Prices Using the Long Short-Term Memory (LSTM) Algorithm

LAILAN SAHRINA HASIBUAN^{1*}, YANDA NOVIALDI¹

Abstrak

Kenaikan harga bahan kebutuhan pokok yang sangat signifikan akan menimbulkan dampak negatif bagi perekonomian masyarakat Indonesia, salah satunya adalah penurunan daya beli. Berdasarkan monitor Pusat Informasi Harga Pangan Strategis dari November 2021 hingga Agustus 2022, minyak goreng merupakan salah satu bahan pokok yang mengalami kenaikan harga yang sangat signifikan. Kenaikan ini tersebar merata di 34 provinsi Indonesia, termasuk provinsi Jawa Barat. Kenaikan yang signifikan ini dapat dicegah dengan melakukan tindakan preventif jauh hari sebelumnya, jika kenaikan ini telah diprediksi sebelumnya. *Deep Learning* merupakan metode *supervised learning* yang banyak digunakan saat ini karena keandalannya untuk menyelesaikan berbagai masalah di bidang penambangan data. *Deep learning* dapat melakukan prediksi harga minyak goreng untuk masa yang akan datang menggunakan data deret waktu. Penelitian ini mengembangkan model untuk memprediksi harga minyak goreng berbentuk curah maupun kemasan menggunakan *deep learning* yang khusus mengelola data *time series* yaitu *Long Short-Term Memory* (LSTM). Berdasarkan metrik evaluasi NRMSE, model yang dibangun mampu mengenali pola harga minyak goreng berbentuk curah maupun kemasan. Nilai NRMSE model LSTM pada proses pelatihan adalah 0.019 untuk pelatihan data minyak goreng curah, dan 0.037 untuk data minyak goreng kemasan.

Kata kunci: Data *deep learning*, Data *time series*, Long Short-Term Memory (LSTM), Minyak goreng, Prediksi

Abstract

A significant increase of the price of basic necessities will affect the economy of the Indonesian people, such as lowering purchasing power. Based on the monitoring of the Strategic Food Price Information Center from November 2021 to August 2022, cooking oil is a necessity that experienced a very significant increase of price in Indonesia. This increase was spread evenly across 34 provinces of Indonesia, including the province of West Java. This significant increase can be prevented by taking preventive actions before, if this increase has been predicted. Deep Learning is a supervised learning method that is widely used today

¹Departemen Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Bogor 16680

*Surel: lailan.sahrina@apps.ipb.ac.id

because of its reliability in solving various problems in the field of data mining. Deep learning can predict future cooking oil prices using time series data. This study develops a model to predict the price of cooking oil in bulk and packaged form using deep learning that specifically manages time series data, namely Long Short-Term Memory (LSTM). Based on the NRMSE evaluation metric, the model built is able to recognize the price fluctuation of cooking oil in the form of bulk and packaging. The NRMSE value of the LSTM model in the training process is 0.019 for bulk cooking oil data training, and 0.037 for packaged cooking oil data.

Keywords: Deep Learning Data, Time Series Data, Long Short-Term Memory (LSTM), Cooking Oil, Prediction

PENDAHULUAN

Minyak goreng merupakan kebutuhan pokok bagi masyarakat Indonesia. Dalam kehidupan sehari-hari hampir seluruh masyarakat Indonesia mengonsumsi minyak goreng, baik masyarakat yang tinggal di perkotaan maupun di pedesaan. Berdasarkan data Gabungan Pengusaha Kelapa Sawit Indonesia (GAPKI), konsumsi lokal minyak sawit Indonesia pada tahun 2021 adalah 18.42 juta ton atau 36% dari total produksi.

Berdasarkan data BPS (2022), minyak goreng pada bulan April 2022 merupakan salah satu komoditas yang mengalami kenaikan harga. Indeks Harga Perdagangan Besar (IHPB) Umum Nasional naik 0.97%. Hal ini tentunya merupakan kenaikan harga yang sangat signifikan yang dapat berdampak bagi kehidupan masyarakat khususnya kalangan menengah ke bawah.

Dampak dari perubahan harga yang signifikan dapat diminimalisir jika langkah-langkah pencegahan telah dilakukan sebelumnya. Oleh karena itu, prediksi harga pada masa yang akan datang diperlukan agar tindakan pencegahan dapat dilakukan lebih awal. Fluktuasi harga minyak goreng dari waktu ke waktu dapat digunakan untuk membangun model prediksi harga yang akan datang. Beberapa penelitian tentang peramalan harga minyak goreng telah dilakukan sebelumnya. Susila dan Munadi (2008) memodelkan harga minyak goreng menggunakan metode ARIMA. Fitri *et al.* (2020) memodelkan harga minyak goreng di tengah pandemi covid-19 di kota Bandar Lampung menggunakan metode ARIMA. Adinugroho (2021) melakukan pendekatan *clustering time series* untuk memodelkan harga minyak goreng. Penelitian ini menghasilkan nilai MAPE pada kisaran 1%–18%.

Di antara penelitian yang telah dipaparkan sebelumnya, belum ada prediksi harga minyak goreng menggunakan *deep learning*, suatu model pembelajaran mesin yang banyak digunakan saat ini karena kemampuannya yang tinggi dalam menyelesaikan berbagai masalah seperti regresi dan klasifikasi. Bagian dari *deep learning* yang memiliki kemampuan khusus untuk menangani data *time series* adalah *Long Short-Term Memory (LSTM)* (Zahara *et al.* 2019). Selain itu, metode LSTM juga dikatakan sebagai model pembelajaran mesin komparatif yang lebih akurat dalam memprediksi harga pangan dibandingkan metode sejenis multilayer perceptron (MLP) dan ARIMA (Sen *et al.* 2020).

Pada penelitian ini dikembangkan model LSTM untuk memprediksi harga minyak goreng di Indonesia. Data latih yang digunakan adalah harga minyak goreng curah dan kemasan pada tanggal 01 November 2021 – 31 Agustus 2022. Rentang waktu ini dipilih karena pada rentang tersebut terdapat kenaikan harga minyak goreng yang signifikan.

METODE

Data Penelitian

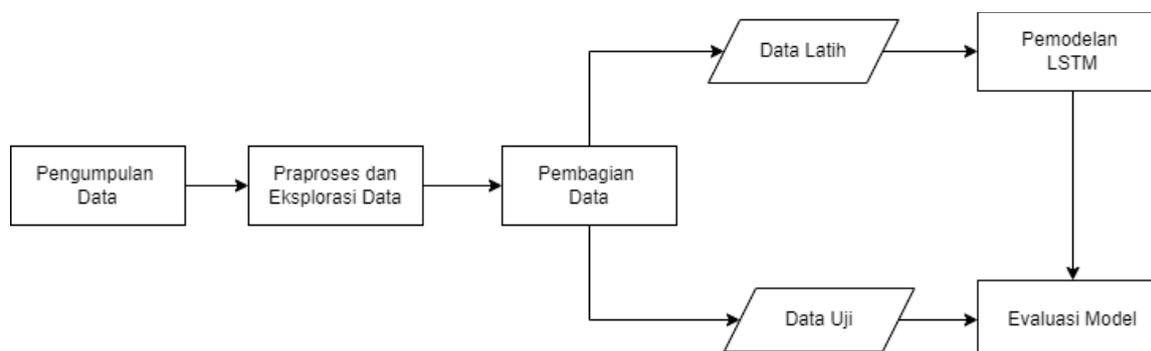
Penelitian ini menggunakan data harga minyak goreng curah dan kemasan pada pasar tradisional di Propinsi Jawa Barat dari tanggal 01 November 2021 sampai 31 Agustus 2022. Data harga minyak goreng curah dan kemasan disimpan dalam file yang terpisah dengan format *Comma Separated Values* (CSV). Data ini masing-masing terdiri dari dua atribut yaitu tanggal dan harga minyak goreng (curah/kemasan). Tabel 1 menyajikan informasi data minyak goreng curah/kemasan yang tersedia.

Tabel 1 Daftar atribut *dataset* komoditas minyak goreng curah dan kemasan

Nama atribut	Jenis data	Keterangan
Tanggal	Numerik, Datetime	Tanggal terjadinya perkembangan harga minyak goreng curah di Jawa Barat
Harga minyak goreng	Time series, Numerik	Harga Komoditas minyak goreng curah pada pasar tradisional dalam satuan kilogram

Tahapan Penelitian

Tahapan penelitian dapat dilihat pada Gambar 1. Tahapan ini terdiri dari beberapa tahapan yaitu pengumpulan data, praproses dan eksplorasi data, pembagian data, pembangunan model prediksi menggunakan LSTM, dan evaluasi model.



Gambar 1 Tahap alur penelitian

Pengumpulan Data

Data harga minyak goreng diunduh melalui halaman web Pusat Informasi Harga Pangan Strategis (PIHPS). Data yang telah diunduh memiliki beberapa tanggal yang hilang dan format yang belum sesuai, oleh sebab itu dilakukan pengolahan data ke dalam file Microsoft Excel secara sekuensial berdasarkan kolom atribut dan disimpan dalam format CSV untuk dilakukan pengolahan data dengan bahasa pemrograman Python.

Praproses dan Eksplorasi Data

Data yang sudah dikumpulkan tidak bisa langsung diproses untuk membuat model. Akan tetapi perlu dilakukan praproses data terlebih dahulu. Data diubah ke dalam format data deret waktu harian. Setelah itu dilakukan pengecekan nilai data yang hilang (*missing value*), seleksi atribut dan normalisasi data. Apabila terdapat *missing value* dilakukan pengisian data yang hilang dengan metode *Last Observation Carried Forward* (LOCF) dengan bantuan fungsi *na.locf* pada pemrograman R. LOCF adalah salah satu teknik imputasi dengan mengisi data

yang tidak lengkap dengan nilai data terakhir sebelum data tidak lengkap. Apabila data tidak lengkap terdapat pada awal data dapat diganti dengan nilai data paling awal data lengkap (Enders, 2010). Pemilihan LOCF dilakukan karena cara kerjanya sesuai dengan kriteria dataset yang digunakan. Fluktuasi harga minyak goreng tidak mengalami perubahan signifikan pada rentang waktu yang dekat. Pada dataset ini, *missing value* terjadi pada rentang waktu yang dekat sehingga imputasi menggunakan data terakhir sebelum *missing* dapat dikatakan sesuai untuk dataset ini.

Pembagian Data

Data yang telah melewati tahap praproses dibagi menjadi dua bagian, yaitu data latih dan data uji. Data latih adalah data yang digunakan untuk melatih algoritme dalam mencari model yang sesuai, sedangkan data uji adalah data yang dipakai untuk menguji dan mengetahui model yang didapat pada saat pengujian data. Menurut Jauhari (2016) dari hasil pengujian menunjukkan bahwa semakin banyak data latih, maka akan menghasilkan akurasi yang lebih akurat. Dalam penelitian ini, persentase data latih yang dipilih sebesar 80% dan data uji sebesar 20%. Nilai data latih yang digunakan pada dataset yaitu dari selang waktu 1 November 2021 sampai 1 Juli 2022, sedangkan nilai data uji yang digunakan yaitu dari selang waktu 2 Juli 2022 sampai 31 Agustus 2022.

Pemodelan Menggunakan LSTM

Sebelum melakukan pemodelan menggunakan LSTM perlu dilakukan inisiasi parameter yang diperlukan LSTM. Pada tahap inisiasi parameter, parameter yang dibutuhkan adalah jumlah *neuron*, *batch size*, jumlah *epoch* dan *optimizer* berupa *learning rate*. Optimizer merupakan salah satu parameter yang dibutuhkan untuk membangun model dalam *package Keras* versi 2.3. Optimizer memiliki peranan penting untuk meningkatkan akurasi sebuah model. Dalam penelitian ini, inisialisasi parameter dilakukan dengan metode *grid search*.

Evaluasi Model

Evaluasi model merupakan tahapan untuk melihat tingkat kebaikan dari model yang telah dibentuk. Evaluasi model dilakukan menggunakan *normalized Root Mean Square Error* (RMSE). *Normalized Root Mean Square Error* (RMSE) adalah indikator dari rata-rata kesalahan (*average error*) pada analisa.

$$NRMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2}}{\frac{1}{N} \sum_{i=1}^N x_i}$$

dengan:

N = banyaknya data

x_i = nilai prediksi saat ke- i

y_i = nilai aktual saat ke- i

HASIL DAN PEMBAHASAN

Praproses dan Eksplorasi Data

Data harga minyak goreng curah dan kemasan yang telah diunduh dari web Pusat Informasi Harga Pangan Strategis (PIHPS), selanjutnya diolah terlebih dahulu menggunakan microsoft excel dengan atribut tanggal dan harga, yang masih-masing disimpan dalam file yang terpisah. Format tanggal dan harga disesuaikan terlebih dahulu agar dapat dibaca dalam program python. Eksplorasi data adalah langkah untuk memahami bagaimana keadaan data sebelum dilakukan praproses. Hal ini dapat digunakan untuk menentukan teknik-teknik praproses yang dibutuhkan. Tabel 2 menunjukkan nilai statistik deskriptif dari setiap atribut sebelum dilakukan imputasi *missing value*, mulai dari nilai rata-rata harga selama periode November 2021 – Agustus 2022, simpangan baku, nilai minimum dan nilai maksimum.

Tabel 2 Statistika deskriptif atribut data harga minyak goreng curah dan kemasan

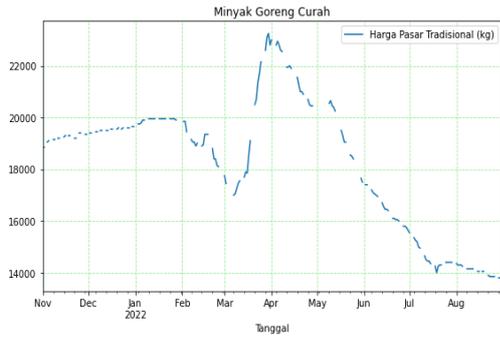
Dataset	Atribut	Nilai minimum	Nilai maksimum	Rata-rata	Simpangan Baku
Komoditas minyak goreng curah	Harga minyak goreng curah	13.750	23.250	18.193	2.565,55288
Komoditas minyak goreng kemasan	Harga minyak goreng kemasan	17.300	25.500	21.585	2.668,18155

Praproses data yang terlebih dahulu dilakukan adalah pengecekan *missing value*. *Missing value* dapat terjadi karena nilai tersebut tidak relevan dengan kasus tertentu, tidak dapat direkam pada saat pengumpulan data atau disebabkan adanya privasi atau perlindungan data (Gimpy *et al.* 2014). Pengecekan *missing value* dilakukan dengan menggunakan fungsi *isnull.sum* pada pemrograman python. *Dataset* komoditas minyak goreng curah dan kemasan di propinsi Jawa Barat yang digunakan pada penelitian ini mengandung *missing value*. Tabel 3 menunjukkan jumlah *missing value* pada atribut harga minyak goreng curah dan kemasan.

Tabel 3 Atribut yang memuat *missing value* pada dataset

Dataset	Atribut	Selang waktu	Missing value	Persentase <i>missing value</i>
Komoditas minyak goreng curah	Harga minyak goreng curah	November 2021 – Agustus 2022	109	35,85 %
Komoditas minyak goreng kemasan	Harga minyak goreng kemasan	November 2021 – Agustus 2022	109	35,85 %

Proses imputasi pada data *missing value* sangat diperlukan karena performa LSTM dalam melakukan prediksi akan sangat berpengaruh jika terdapat *missing value*. Gambar 2-5 menunjukkan grafik sebelum dan sesuai imputasi data *missing value* baik pada data minyak goreng curah maupun kemasan.



Gambar 2 Plot data minyak goreng curah sebelum dilakukan imputasi data



Gambar 3 Plot data minyak goreng kemasan setelah dilakukan imputasi data



Gambar 4 Plot data minyak goreng kemasan sebelum dilakukan imputasi data



Gambar 5 Plot data minyak goreng kemasan setelah dilakukan imputasi data

Nilai atribut harga minyak goreng curah dan kemasan dinormalisasi dengan mengubah skala data. Proses normalisasi dilakukan dengan menggunakan library MinMaxScaler yang mengubah ukuran data dari rentan asli, sehingga semua nilai berada pada rentan 0 sampai 1. Tabel 4 menunjukkan data sebelum dan sesudah dinormalisasi.

Tabel 4 Contoh data harga minyak goreng curah dan kemasan setelah dinormalisasi

Dataset	Tanggal	Data sebelum dinormalisasi	Data setelah dinormalisasi
Harga minyak goreng curah	2021-11-01	18850	0.5368421
	2021-11-02	18550	0.5368421
Harga minyak goreng kemasan	2021-11-01	17800	0.06097579
	2021-11-05	18000	0.08536601

Pemodelan Prediksi Menggunakan Long Short-Term Memory (LSTM)

Data yang digunakan untuk melakukan pemodelan prediksi adalah data harga minyak goreng curah dan kemasan di Jawa Barat. Data dibagi menjadi 80% data latih dan 20% data uji. Model dibangun menggunakan metode LSTM yang ada di library keras pada pemrograman Python. Arsitektur LSTM untuk prediksi harga minyak goreng curah dan kemasan pada penelitian ini ditunjukkan pada tabel 5.

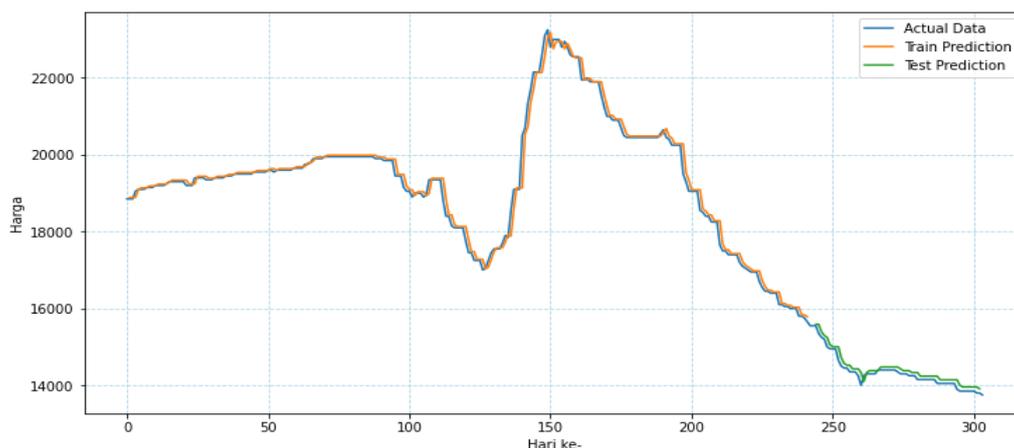
Tabel 5 Arsitektur dan Parameter Model Prediksi Menggunakan LSTM

Karakteristik	Spesifikasi
Arsitektur	3 LSTM layer dengan 50, 500 unit neuron 1 Dense layer
Optimizer	Adam
Batch size	32, 64
Epoch	100, 1000
Learning rate	0.01, 0.001, 0.005

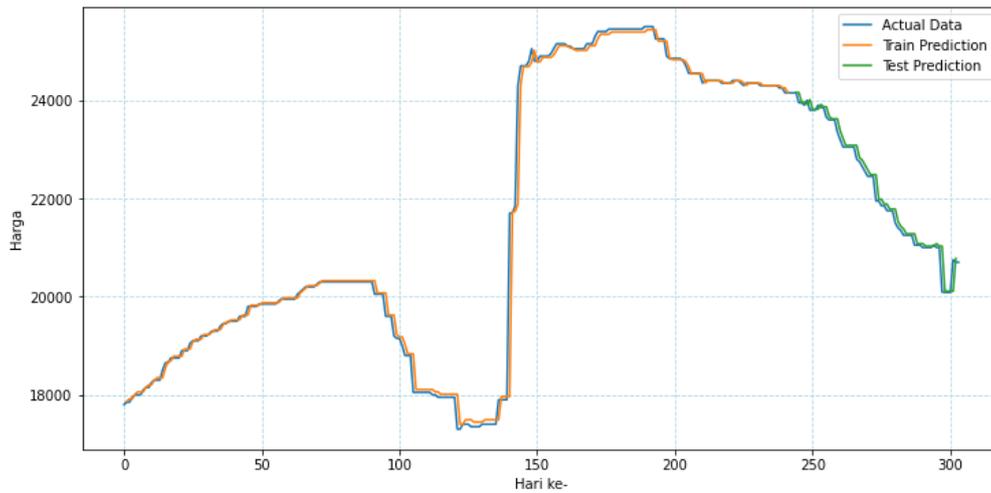
Penelitian ini menggunakan 3 LSTM layer dan 1 Dense layer. Jumlah neuron yang dicobakan yaitu 50 dan 100. Sebenarnya tidak ada struktur referensi khusus untuk memilih jumlah neuron di setiap layer LSTM (Le *et al.* 2019). Namun hal ini dilakukan untuk mengetahui berapa banyak neuron yang ideal untuk pembuatan model deret waktu prediksi harga minyak goreng curah dan kemasan. Optimizer yang digunakan dalam arsitektur LSTM adalah Adaptive Moment Estimation (Adam), *batch size* yang digunakan sebanyak 32 dan 64, dengan *epoch* sebanyak 100 dan 1000 sedangkan untuk learning rate yaitu sebesar 0.01; 0.001; dan 0.005.

Evaluasi Model

Prediksi Menggunakan *Long Short-Term Memory* (LSTM) dilakukan sepanjang periode November 2021 hingga Agustus 2022 berdasarkan persentase data latih dan data uji. Evaluasi model dilakukan dengan menghitung nilai *normalized Root Mean Square Error* (NRMSE). Berdasarkan percobaan yang telah dilakukan didapat model terbaik untuk prediksi harga minyak goreng curah dan kemasan dengan jumlah *neuron* 50, *batch size* 64, *epoch* 1000 dan *learning rate* 0.01. Nilai NRMSE minyak goreng curah sebesar 0.019 dan nilai NRMSE Minyak goreng kemasan sebesar 0.037. Gambar 6 dan 7 menunjukkan perbandingan antara nilai aktual (biru) data harga minyak goreng, nilai prediksi data latih(jingga), dan nilai prediksi data uji (hijau) menggunakan LSTM. Berdasarkan gambar tersebut dapat dilihat bahwa hasil prediksi mampu mengikuti pola yang dibentuk data *dataset training* serta mempelajari pola-pola kejadian terhadap data *time series*.



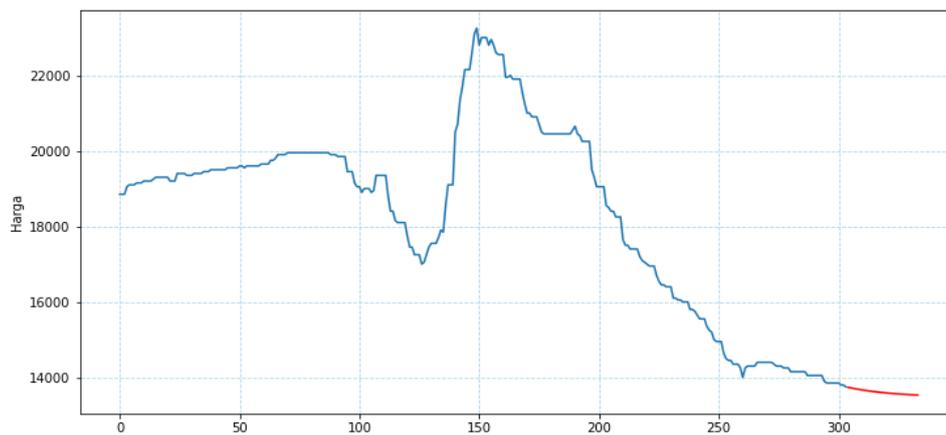
Gambar 6 Plot perbandingan data aktual dengan hasil prediksi menggunakan data training dan data uji minyak goreng curah dengan nilai *neuron* 50, *batch size* 64, *epoch* 1000 dan *learning rate* 0.01.



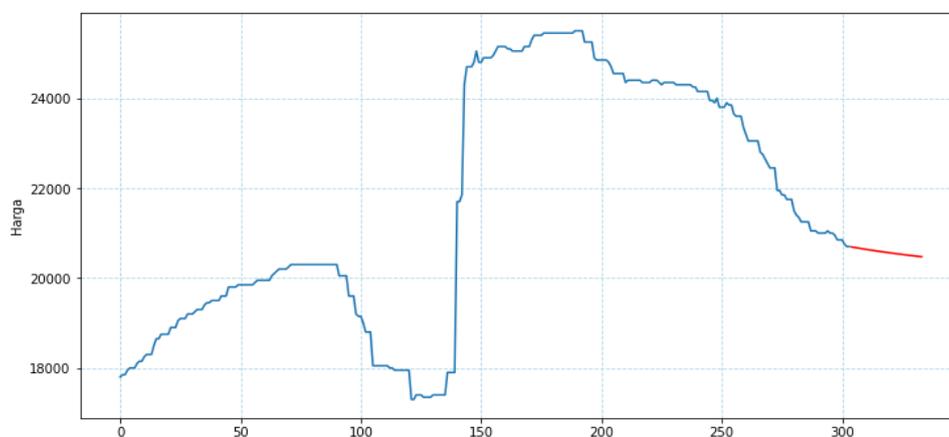
Gambar 7 Plot perbandingan data aktual dengan hasil prediksi menggunakan data training dan data uji minyak goreng kemasan dengan nilai *neuron* 50, *batch size* 64, *epoch* 1000 dan *learning rate* 0.01.

Hasil Prediksi Menggunakan *Long Short-Term Memory (LSTM)*

Hasil prediksi harga minyak goreng curah dapat dilihat pada Gambar 8. Sedangkan hasil prediksi harga minyak goreng kemasan dapat dilihat pada Gambar 9.



Gambar 10 Plot hasil prediksi harga minyak goreng curah 01- 30 September 2022



Gambar 11 Plot hasil prediksi harga minyak goreng kemasan 01- 30 September 2022

SIMPULAN

Pemodelan prediksi harga minyak goreng curah dan kemasan menggunakan algoritme *long short-term memory* (LSTM) dapat menerapkan metode imputasi data time series menggunakan imputasi *Last Observation Carried Forward* (LOCF). Optimasi parameter dengan menggunakan 80% data latih dan 20% data validasi berhasil mendapatkan model LSTM yang optimal berdasarkan nilai NRMSE yang kecil yaitu 0.019 untuk minyak goreng curah dan 0.039 untuk minyak goreng kemasan. Optimasi ini menggunakan *neuron* sebanyak 50, *learning rate* 0.01, *batch size* 64 dan *epoch* 1000. Model ini dapat memprediksi perubahan harga minyak goreng yang signifikan di pasar tradisional Jawa Barat.

DAFTAR PUSTAKA

- [BPS] Badan Pusat Statistik. Indeks Harga Perdagangan Besar (IHPB) Umum Nasional naik 0,97 persen. Jakarta (ID): Badan Pusat Statistik. 2022
- Adinugroho W. 2021. Pendekatan Clustering time series pada Peramalan Harga Minyak Goreng. *Jurnal Ilmiah Populer*. 4: 47-55
- Enders, 2010. CK Enders. Applied missing data analysis. Guilford Press, New York (2010)
- Fitri A, Margasaty F, Kusmaria, Desfaryani R, Dewi VU. 2020. Peramalan Harga Minyak Goreng di Tengah Pandemi Covid-19 Kota Bandar Lampung, 10(1): 21-26
- Gimpy, Vohra R, Minakshi. 2014. Estimation of Missing Values Using Decision Tree Approach. *International Journal of Computer Science and Information Technologies*, Vol. 5 (4): 5216-5220
- Jauhari D, Himawan A, Dewi C. 2016. Prediksi Distribusi Air PDAM Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation di PDAM Kota Malang. *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*. 3(2): 83-87
- Le XH, Ho HV, Lee G, Jung S. 2019. Application of Long Short-Term Memory (LSTM) neural network for flood forecasting. *Water (Switzerland)*. 11(7). doi:10.3390/w11071387
- Sen S, Sugiarto D, Rochman A. 2020. Komparasi Metode Multilayer Perceptron (MLP) dan Long Short Term Memory (LSTM) dalam Peramalan Harga Beras. *Jurnal ULTIMATICS*. 12(1): 39-40
- Susila WR, Munadi E. 2008. Peramalan Harga Eceran Minyak Goreng dengan Model ARIMA. *Jurnal Ekonomi*. 3(7): 40-53
- Zahara S, Sugianto, Ilmiddafiq MB. 2019. Prediksi Indeks Harga Konsumen Menggunakan Long Short Term Memory (LSTM) Berbasis Cloud Computing. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*. 3(3): 357-363