

PERBANDINGAN MODEL LSTM DAN GRU UNTUK MEMPREDIKSI HARGA MINYAK GORENG DI INDONESIA

Mochammad Agus Sholeh^{1*}, Rahmat Hidayat²

^{1,2}Universitas Singaperbangsa Karawang, Indonesia

*e-mail: mochammad.agus18098@student.unsika.ac.id

Abstract: Cooking oil, a food ingredient used for cooking, has increased in price in Indonesia. Based on the Indonesian Strategic Food Price Center data, cooking oil reached twice the regular price at the beginning of 2022. Long Short-Term Memory and Gated Recurrent Units are several ways to predict time domain data. The data set for prediction is cooking oil price data for the last three years in the Indonesian Strategic Food Price Center data. The data taken is pre-processed by changing the data type and filling in the blank day data with the previous day's value. Data that has been pre-processed is modeled using Long Short-Term Memory and Gated Recurrent Units. Modeling obtained Error, Loss, and selection of the best unit and dropout model parameters in making predictions for both types of models. As a result, this research compared the prediction results of the Long Short-Term Memory model and the Gated Recurrent Unit on the cooking oil data set and predict price changes that will occur during the following year.

Keywords: *Cooking Oil, Gated Recurrent Unit, Long Short-Term Memory*

Abstrak: Minyak goreng merupakan bahan pangan yang digunakan sebagai bahan untuk memasak mengalami kenaikan harga di Indonesia. Tercatat pada data Pusat Harga Pangan Strategis Indonesia, minyak goreng mencapai harga dua kali lipat dari harga normal pada awal tahun 2022. Long Short-Term Memory serta Gated Recurrent Unit merupakan beberapa cara untuk memprediksi data domain waktu. Data set untuk prediksi merupakan data harga minyak goreng selama tiga tahun terakhir pada data Pusat Harga Pangan Strategis Indonesia. Data yang telah diambil dilakukan pre-processing dengan mengubah tipe data dan mengisi data hari yang kosong dengan nilai hari sebelumnya. Data yang telah dilakukan pre-processing dilakukan pemodelan dengan Long Short-Term Memory serta Gated Recurrent Unit. Pemodelan didapatkan Error, Loss dan pemilihan parameter model unit dan dropout terbaik dalam melakukan prediksi untuk kedua tipe model. Hasil penelitian ini menganalisis perbandingan hasil prediksi model Long Short-Term Memory beserta Gated Recurrent Unit pada data set minyak goreng dan memprediksi perubahan harga yang akan terjadi selama satu tahun berikutnya.

Kata Kunci: *Gated Recurrent Unit, Long Short-Term Memory, Minyak Goreng*

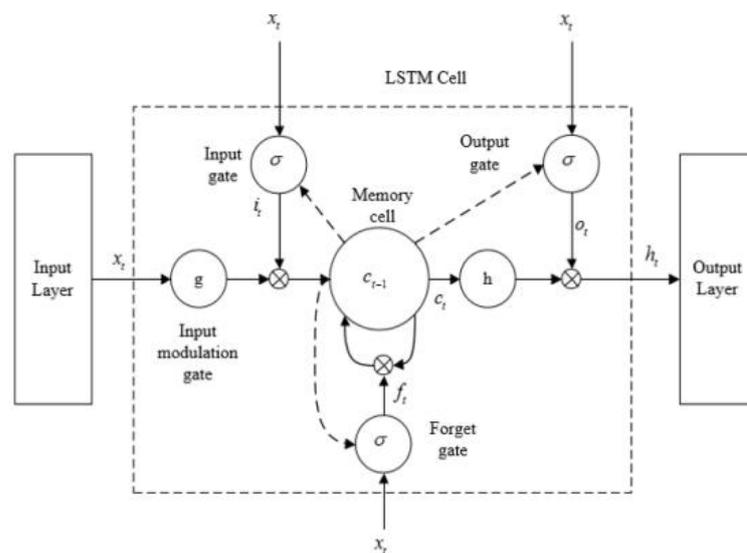
Copyright (c) 2022 The Authors. This is an open access article under the CC BY-SA 4.0 license (<https://creativecommons.org/licenses/by-sa/4.0/>)

PENDAHULUAN

Minyak Goreng merupakan minyak yang dihasilkan oleh pemurnian bagian tumbuhan, hewan, bahkan dibuat secara sintetik untuk digunakan sebagai penggoreng makanan (Sitepoe, Nusantara, & Andarnuswari, 2008). Dari jenisnya, Kurniadi (2022) menyebutkan bahwa minyak goreng yang terdapat di Indonesia

terdiri dari dua tipe, minyak goreng curah yang dijual kepada konsumen tanpa merek, dan minyak goreng bermerek. Minyak goreng yang merupakan bahan untuk memasak mengalami kenaikan harga lebih dari 100 persen harga biasanya pada beberapa wilayah di Indonesia pada awal tahun 2022 (Kompas, 2022). Data Pusat Informasi Harga Pangan Strategis Indonesia (disingkat menjadi PIHPS) mencatumkan temuan harga tertinggi minyak goreng berada pada harga \pm Rp.24.000,00 pada bulan April 2022 (PIHPS, 2020).

LSTM dan GRU merupakan modifikasi arsitektur dari *Recurrent Neural Network* (RNN) yang memiliki *forget gate* namun pada GRU tidak memiliki *output gate* (Britz, 2015; Gers, Schmidhuber, & Cummins, 2000). LSTM dapat memiliki kemampuan mempelajari data, memprediksi suatu data dan banyak digunakan dalam mengelola teks, video, serta data deret waktu tertentu. Di LSTM terdapat istilah *cell state* yang akan diteruskan ke setiap *cell* (Wisyalidin, Luciana, & Pariaman, 2020).



Gambar 1. Sel LSTM

Dalam menjelaskan gambar 1 tersebut, penulis merujuk Wisyalidin, Luciana, & Pariaman (2020) yang menerangkan bahwa LSTM memiliki tiga gerbang, *forget*, *input*, dan *output*, gerbang tersebut berfungsi untuk menghapus ataupun menambah informasi. Adapun yang dimaksud dengan *Forget gate* berfungsi untuk menghapus data dan fitur ini akan digunakan untuk proses *input* setelahnya. Selain itu, *input gate* merupakan titik masuk data dari luar dan melakukan pemrosesan

data masuk, dan output gate untuk memproses perhitungan masuk dan menghasilkan luaran dalam LSTM cell (Aldi, Jondri, & Aditsania, 2018).

Selanjutnya, Aldi, Jondri, & Aditsania (2018) menyebutkan bahwa Model LSTM yang mereka gunakan adalah untuk memprediksi harga Bitcoin yang menghasilkan model dengan akurasi bernilai 95.36%. Chimmula & Zhang (2020) membuat model LSTM yang berbeda dengan fungsinya digunakan untuk memprediksi transmisi COVID-19 di Kanada.

Serupa dengan LSTM adalah Gated Recurrent Unit (GRU) merupakan jenis modifikasi arsitektur dari RNN (Cho, Van Merriënboer, Bahdanau, & Bengio, 2014). Britz (2015) & Gers dan Schmidhuber, & Cummins (2000) memandang bahwa GRU seperti LSTM memiliki *Forget Gate*. Namun begitu, Gruber, & Jockisch (2020) dan Chung, Gulcehre, Cho, & Bengio (2014) berpandangan bahwa *Forget Gate* memiliki parameter yang lebih sedikit dari LSTM karena tidak ada *Output Gate*. Dengan kata lain bahwa GRU memiliki performa yang lebih baik pada data set yang lebih sedikit.

Pusat Informasi Harga Pangan Strategis Nasional (PIHPS) merupakan sebuah *website* negara yang memberikan informasi terkait harga bahan pangan seperti Minyak Goreng, Beras Daging, Cabai, dan bahan pangan lainnya. PIHPS (2022) menjelaskan bahwa PHPS sendiri merupakan *database* harga pangan seluruh Indonesia. Minyak goreng yang memiliki perubahan harga lebih dari 100 persen yang menjadi keluhan masyarakat Indonesia menjadi dasar dari tujuan penelitian ini.

Dari hasil obesrvasi tipe model, model LSTM dan GRU merupakan tipe model untuk data seri waktu yang akan digunakan. Penelitian ini mengimplementasi, menganalisis serta membandingkan hasil dari kedua prediksi model LSTM serta GRU terhadap data set PIHPS untuk harga minyak goreng di Indonesia selama satu tahun berikutnya. Tujuan penelitian ini untuk mendapatkan model yang memiliki performa terbaik untuk mengetahui perkembangan harga minyak goreng yang saat ini sangat dibutuhkan untuk diketahui oleh masyarakat Indonesia.

METODE

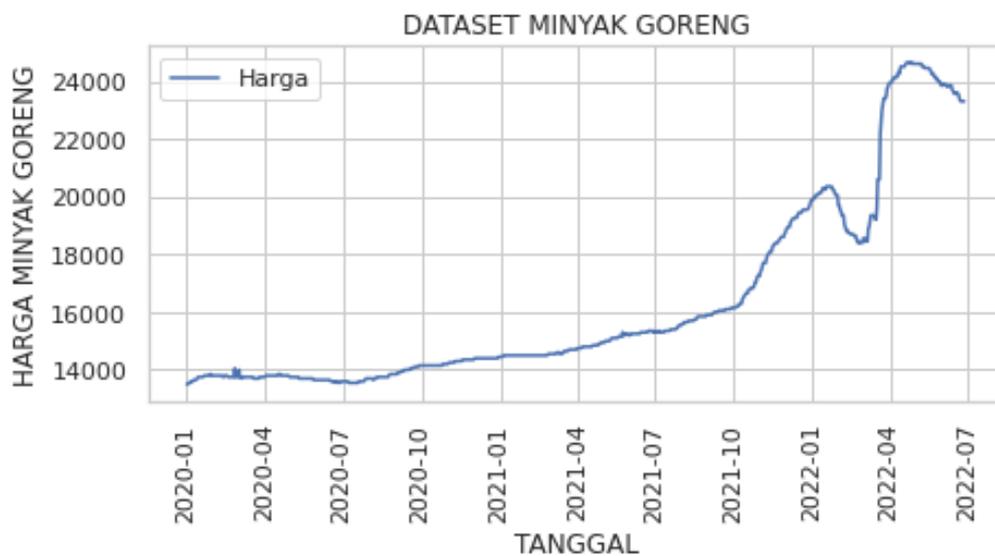
Pengumpulan Data Set Harga Minyak Goreng

Data set harga minyak goreng bersumber dari aplikasi berbasis situs web milik pemerintah yaitu PIHPS Nasional. Data set minyak goreng diambil dari seluruh provinsi di Indonesia. Digunakannya data set harga minyak goreng dengan waktu harian serta diambil data set tiga tahun terakhir.

Pre-processing Data Set Prediksi Harga Minyak Goreng

Pemrosesan data awal bertujuan untuk mengubah tipe data objek atau string menjadi float, dan mengisi nilai serta hari yang kosong. Alat yang digunakan untuk memproses data awal meliputi: Pandas, Numpy, dan Google Colaboratory. Pemrosesan data awal dimulai dengan mengekspor data dari PIHPS dalam bentuk tabular (.xlsx), lalu mengimpor pustaka dan membuat variabel serta data frame. Data frame yang digunakan meliputi baris yang berisi tanggal dan nilai harga minyak per hari.

Transposisi data frame dilakukan sehingga baris dan kolom ditukar untuk keperluan pemodelan. Fungsi *Regex* dipanggil untuk memangkas huruf yang terdapat pada kolom nilai sehingga hanya tersisa angka saja. kemudian mengubah tipe data menjadi *float*. Kolom tanggal diubah menjadi bentuk tanggal dengan mengimpor pustaka *datetime*, dilakukan pengisian hari yang hilang. Selanjutnya memanggil fungsi *ffill* untuk mengisi nilai-nilai yang kosong.



Gambar 2. Visualisasi data set harga minyak goreng

Pada Gambar 2. Menunjukkan hasil visualisasi data harga minyak goreng di Indonesia selama tiga tahun terakhir yang telah di *pre-process* dimana kenaikan tertinggi terjadi untuk akhir bulan Maret 2022 sampai awal bulan April 2022. Data yang telah di *pre-process* akan dilakukan pemodelan dengan menggunakan model LSTM dan GRU.

Pre-processing data selanjutnya yaitu membagi data *training* dan data *testing*. Dalam pembagian data set untuk pemodelan baik menggunakan LSTM maupun GRU digunakan 75% sebagai data *training* dan 25% sebagai data *testing*. Berikut Tabel rincian pembagian data setnya.

Tabel 1. Pembagian Data Set Model LSTM dan GRU

Model	Training Data	Testing Data	Total Data
LSTM	682	227	909
GRU	682	227	909

Dilakukan proses *feature scaling* menggunakan *MinMaxScaler* agar dapat digunakan kembali pada data tes maupun data observasi yang baru. Hasil *scaling* yang dibuat disimpan pada kolom baru yaitu '*scaled*'. Fungsi *sliding window* digunakan pada data dengan *window size* sebesar 30 dan diterapkan pada data *train* dan data *test* yang telah di *scaling*.

Pemodelan Model LSTM dan GRU untuk Prediksi Harga Minyak Goreng

Pemodelan dilakukan dengan mengimpor pustaka *Tensorflow Keras*. Fungsi model yang dibuat memiliki input *shape* 30,1 , *Dense* layer 32 neuron dengan *activation* function *ReLu*, *dropout* diantara *Dense* layer dan *Dense* output layer dengan *dense* output layer berjumlah satu neuron. *Metric* yang digunakan *Mean Absolute Error* (MAE), model *Loss function* dengan *Mean Squared Error* (MSE) dengan *optimizer* Adam. Untuk model LSTM dan GRU memiliki unit yang sama berjumlah 64 dan persentase *dropout* yang sama berjumlah 0.2 dan *epoch* yang akan dilakukan sebesar 100.

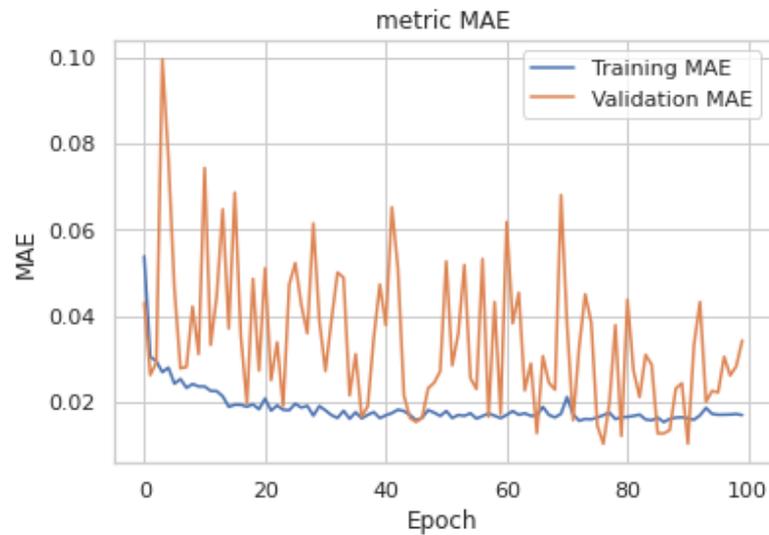
Hyper tuning Parameter dilakukan pada kedua model dengan menambahkan kombinasi nilai unit berjumlah 16, 32, 64, 128 dan *dropout* 0.1 dan 0.2. Variabel *GridSearch* dibuat dengan memasukkan beberapa parameter seperti *estimator* fungsi model yang akan dilakukan *GridSearch*, parameter yang akan diuji seperti

unit dan *dropout*, jumlah pekerjaan untuk dikerjakan secara paralel, dan banyaknya *k-fold cross validation* yang berjumlah 5.

HASIL DAN PEMBAHASAN

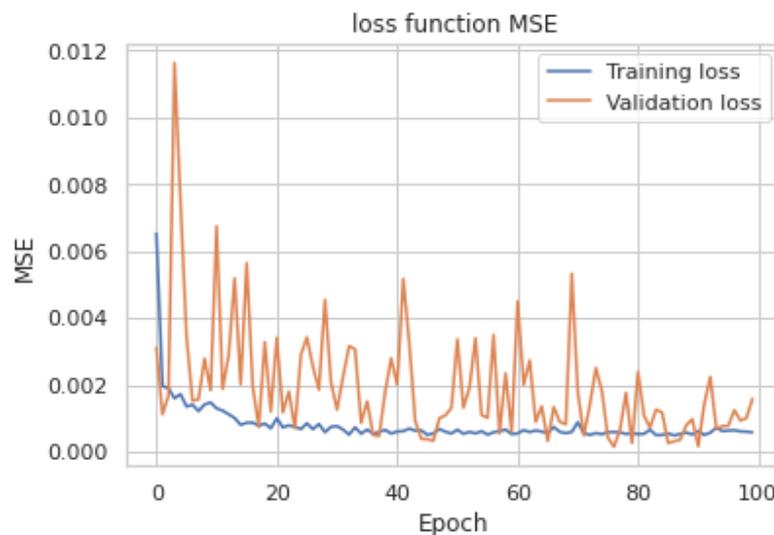
Hasil

Hasil Prediksi Model LSTM



Gambar 3. Mean Absolute Error Long Short-Term Memory Model

Didapatkan pada Gambar 3. hasil pemodelan menggunakan LSTM dengan *epoch* 100, *Mean Absolute Error* (MAE) *validation* tertinggi pada nilai ± 0.10 dengan *epoch* kurang dari 20. Pada *epoch* 80-100 didapatkan hasil MAE *validation* tertinggi bernilai ± 0.04 . MAE training didapatkan ± 0.019 .



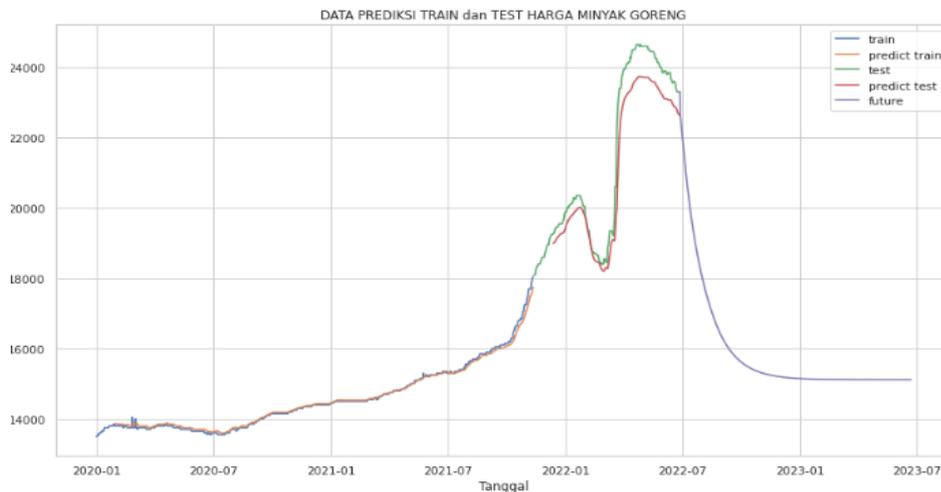
Gambar 4. Mean Squared Error Long Short-Term Memory Model

Pada Gambar 4. hasil pemodelan menggunakan LSTM dengan *epoch* 100, Mean Squared Error (MSE) *Validation loss* tertinggi pada nilai ± 0.012 dengan *epoch* kurang dari 20. Pada epoch 80-100 didapatkan hasil MSE *Validation loss* tertinggi bernilai ± 0.002 . MSE training *loss* didapatkan ± 0.0009 .

Tabel 2. Hasil Parameter Model LSTM

LSTM unit	Dropout	Mean	Standard Deviation
16	0.1	-0.000122	0.000098
16	0.2	-0.000210	0.000233
32	0.1	-0.000599	0.000908
32	0.2	-0.000126	0.000059
64	0.1	-0.000164	0.000144
64	0.2	-0.000148	0.000216
128	0.1	-0.000085	0.000078
128	0.2	-0.000071	0.000034

Pada Tabel 2. menjelaskan hasil *hyper tuning* unit dan *dropout* pemodelan menggunakan LSTM. Pada model yang dibuat didapatkan *best estimator model* dengan unit 128 dan *dropout* 0.2 yang memiliki *Mean* -0.000071 dan *Standard Deviation* 0.000034.

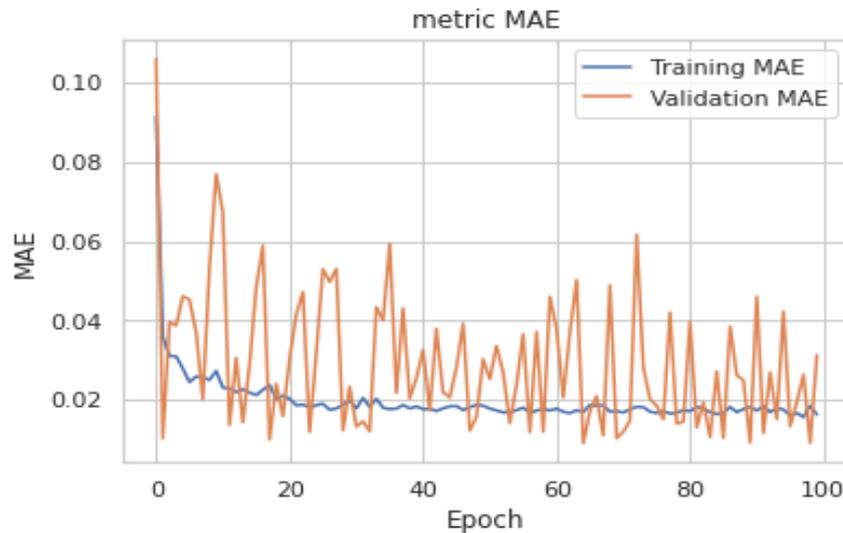


Gambar 5. Hasil prediksi harga minyak goreng satu tahun berikutnya

Pada Gambar 5. menjelaskan prediksi perubahan harga minyak goreng satu tahun berikutnya dengan model LSTM. Dimana pada Maret hingga April 2022 mengalami kenaikan yang tinggi, namun pada bulan Mei 2022 dan seterusnya mengalami penurunan dan diprediksi akan terus turun pada rentang harga

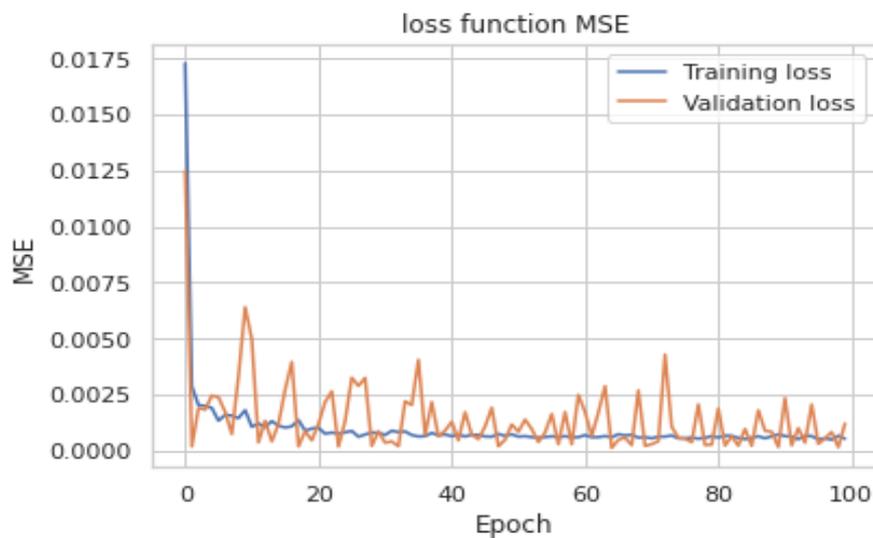
\pm Rp.22.000,00 - \pm Rp.15.000,00 dan akan stabil di tahun 2023 dengan harga berkisar \pm Rp.15.000,00.

Hasil Prediksi Model GRU



Gambar 6. Mean Absolute Error Gated Recurrent Unit Model

Didapatkan pada Gambar 6. hasil pemodelan menggunakan GRU dengan *epoch* 100, *Mean Absolute Error* (MAE) *validation* tertinggi pada nilai \pm 0.15 dengan *epoch* kurang dari 20. Pada *epoch* 80-100 didapatkan hasil MAE *validation* tertinggi bernilai \pm 0.045. untuk MAE training didapatkan \pm 0.019.



Gambar 7. Mean Squared Error Gated Recurrent Unit Model

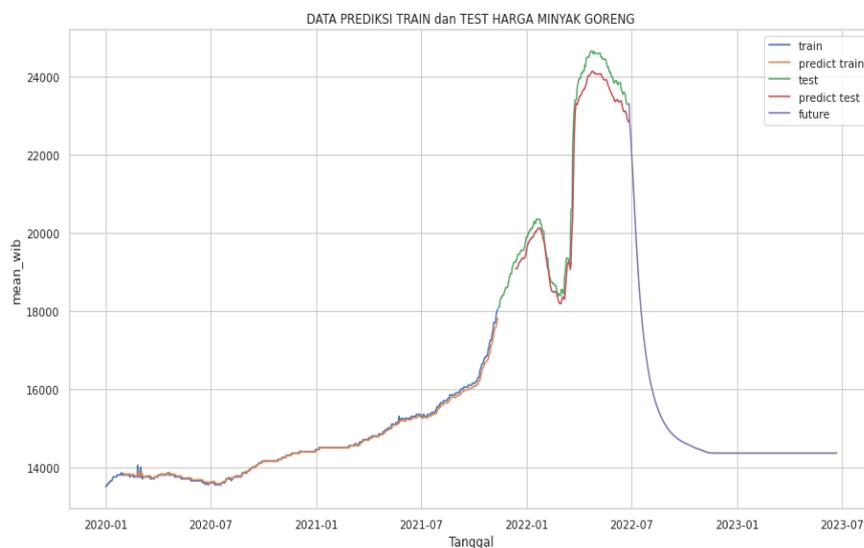
Pada Gambar 7. hasil pemodelan menggunakan GRU dengan *epoch* 100, *Mean Squared Error* (MSE) *Validation loss* tertinggi berada pada nilai \pm 0.0125

dengan *epoch* kurang dari 20. Pada *epoch* 80-100 didapatkan hasil MSE *Validation loss* tertinggi bernilai ± 0.0025 . MSE *training loss* didapatkan ± 0.001 .

Tabel 3. Hasil Parameter Model GRU

GRU unit	Dropout	Mean	Standard Deviation
16	0.1	-0.000664	0.001077
16	0.2	-0.000215	0.000305
32	0.1	-0.000098	0.000061
32	0.2	-0.000335	0.000403
64	0.1	-0.000091	0.000058
64	0.2	-0.000186	0.000150
128	0.1	-0.000207	0.000220
128	0.2	-0.000085	0.000038

Pada Tabel 3. menjelaskan hasil *hyper tuning* unit dan *dropout* pemodelan menggunakan GRU. Pada model yang dibuat didapatkan *best estimator model* dengan unit 128 dan *dropout* 0.2 yang memiliki *Mean* -0.000085 dan *Standard Deviation* 0.000038.



Gambar 8. Hasil prediksi harga minyak goreng satu tahun berikutnya

Gambar 8. menjelaskan prediksi perubahan harga minyak goreng satu tahun berikutnya dengan model LSTM. Dimana pada Maret hingga April 2022 mengalami kenaikan yang tinggi, namun pada bulan Mei 2022 dan seterusnya mengalami penurunan dan diprediksi akan terus turun pada rentang harga

±Rp.22.000,00 hingga ±Rp.14.500,00 dan akan stabil di tahun 2023 dengan harga berkisar ±Rp.14.500,00.

Pembahasan

Berdasarkan hasil penelitian tipe model LSTM dan GRU didapatkan pemilihan parameter terbaik pada dropout dan jumlah unit. dipilih parameter dengan jumlah unit sebesar 128 dan *dropout* 0.2 terbaik. Didapatkan bahwa model LSTM memiliki hasil yang lebih baik dengan mean *Mean* -0.000071 dan *Standard Deviation* 0.000034. Dropout pada model digunakan untuk regularisasi untuk dapat mengurangi *overfitting* serta memperbaiki performa dari model. Untuk unit merupakan jumlah dari *hidden state* dari state output untuk cell RNN.

Model LSTM dan GRU memiliki loss function MAE dan MSE yang digunakan untuk menghitung akurasi dari nilai sebenarnya dan nilai hasil prediksi. MAE merupakan perhitungan error antar pasangan dan observasi yang mengekspresikan fenomena yang sama. MSE merupakan perbandingan nilai prediksi dan nilai sebenarnya. Pada penelitian model LSTM dan GRU didapatkan model terbaik untuk memprediksi harga minyak goreng, Model LSTM memiliki performa yang lebih baik dibandingkan model GRU dengan MAE *validation* bernilai ±0.04 dan MSE *validation loss* bernilai ±0.002.

Untuk MAE dan MSE pada kedua model memiliki nilai yang hampir sama akan tetapi *Error* dan *Loss* pada Model GRU diamati sedikit lebih besar daripada model LSTM. Pada proses prediksi harga minyak goreng untuk satu tahun berikutnya didapatkan bahwa kedua model memprediksi penurunan pada harga minyak goreng dengan model LSTM lebih tinggi dengan harga stabil pada ±Rp.15.000,00 dan prediksi dengan model GRU lebih rendah dan stabil pada ±Rp.14.500,00

SIMPULAN

Model LSTM dan Model GRU dapat memprediksi data domain waktu seperti pada penelitian ini yang menggunakan data set harga minyak goreng yang didapatkan pada situs web pemerintah PIHPS. Model LSTM dan Model GRU dalam MSE dan MAE memiliki nilai *Error* dan *Loss* yang hampir sama namun pada model GRU *Error* dan *Loss* yang didapatkan sedikit lebih besar. Model

LSTM dan Model GRU dengan *tuning hyperparameter* yang sama (jumlah unit dan *dropout*) memiliki parameter terbaik yang sama akan tetapi dengan nilai *mean* dan standar deviasi yang berbeda dengan model GRU yang sedikit lebih besar dalam standar deviasi dan rata-rata yang lebih menjauhi nol daripada model LSTM.

DAFTAR RUJUKAN

- Aldi, M. W. P., Jondri, J., & Aditsania, A. (2018). Analisis dan Implementasi Long Short Term Memory Neural Network untuk Prediksi Harga Bitcoin. *eProceedings of Engineering*, 5(2).
- Britz, D. (2015). Recurrent neural network tutorial, part 4 implementing a gru/lstm rnn with python and theano. URL <http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-grulstm-rnn-with-python-and-theano>.
- Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. *Chaos, Solitons & Fractals*, 135, 109864.
- Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*.
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451-2471. DOI: 10.1162/089976600300015015.
- Gruber, N., & Jockisch, A. (2020). Are GRU cells more specific and LSTM cells more sensitive in motive classification of text? *Frontiers in artificial intelligence*, 3, 40.
- Kurniadi, M. (2022). *Arti Minyak Goreng Curah di Kamus Besar Bahasa Indonesia (KBBI)*".
- Natallia, Fransisca. (2022). Minyak goreng penting trending di Twitter, warganet luapkan kekecewaan." <https://www.kompas.tv/article/271244/minyak->

goreng-penting-trending-di-twitter-warganet-luapkan-kekecewaan#:~:text=JAKARTA%2C%20KOMPAS.TV%20%E2%80%93%20Minyak%20Goreng%20Penting%20menjadi%20trending,50.000%20sedangkan%20harga%20Rp%2023.000%20untuk%20satu%20liter.

PIHPS. (2022). Minyak goreng. <https://hargapangan.id/tabel-harga/pasar-tradisional/daerah>.

Sitepoe, A. Nusantara and D. Andarnuswari. (2022). *Corat-coret anak desa berprofesi ganda*. Jakarta: Kepustakaan Populer Gramedia (KPG).

Wisyalidin, M. K., Luciana, G. M., & Pariaman, H. (2020). Pendekatan LSTM untuk Memprediksi Kondisi Motor 10 kV pada PLTU Batubara. *KILAT*, 9(2), 311-318.