

Prediksi Harga Minyak Dunia Dengan Metode Deep Learning

Muhammad Hussein ¹⁾*, Yufis Azhar ²⁾

Universitas Muhammadiyah Malang ^{1),2)}
hussein.muhammad47@gmail.com ¹⁾*, yufis@umm.ac.id ²⁾

Abstrak

Peramalan seri waktu mendapatkan banyak perhatian dari berbagai penelitian. Salah satu data seri waktu yang berubah setiap periode tertentu adalah minyak bumi. Secara umum harga minyak bumi dipengaruhi oleh dua hal yaitu permintaan dan pendapatan. Pada penelitian ini menggunakan state-of-the-art model Deep Learning LSTM (Long Short Term Memory) untuk meramalkan harga minyak dalam periode tertentu. Metode ini digunakan karena arsitekturnya dapat beradaptasi dengan belajar non-linear dari data seri waktu yang kompleks. Dataset yang digunakan adalah data Brent Oil Price yang selalu di update setiap minggu. Dataset ini berisi harga minyak brent dari tahun 1987 sampai sekarang. Beberapa model yang dibangun terbukti dapat meramalkan harga minyak dengan baik. Model terbaik yang didapatkan dari penelitian ini memiliki RMSE 0,0186 dan MAE 0,013.

Kata kunci: LSTM, deep learning, peramalan, harga minyak

Abstract

[Forecasting World Oil Price with Deep Learning Method] Time series forecasting gets a lot of attention from various studies. One of the time-series data that changes every certain period is petroleum. In general, the price of petroleum is affected by two things, namely demand and income. This research uses a state-of-the-art Deep Learning LSTM (Long Short-Term Memory) model to predict the oil price in a certain period. This method is used because the architecture can adapt to non-linear learning from complex time series data. The dataset used is the Brent Oil Price data, which is always updated every week. This dataset contains the price of Brent oil from 1987 to the present. The models that were built proved to be able to predict oil prices well. The best models obtained from this study have RMSE 0.0186 and MAE 0.013.

Keywords: LSTM, deep learning, forecasting, oil price

1. PENDAHULUAN

Minyak bumi atau petroleum merupakan kebutuhan yang sangat penting dalam memenuhi konsumsi energi dunia. Setiap negara pasti membutuhkan pasokan minyak bumi untuk pemenuhan kebutuhan mereka dari segi ekonomi, militer, pertanian, dan lain-lain. Bagi negara yang bukan penghasil minyak atau produksi minyaknya sedikit akan terus melakukan impor untuk memenuhi kebutuhan energi mereka.

Fluktuasi harga minyak selalu dianggap sebagai barometer ekonomi di seluruh dunia, sehingga setiap perubahan akan harga minyak selalu menjadi isu panas untuk dibahas dalam lingkaran politik dan ekonomi di setiap negara. Fluktuasi harga minyak di pasar internasional pada prinsipnya mengikuti aksioma yang berlaku umum dalam ekonomi pasar, dimana tingkat harga yang berlaku sangat ditentukan oleh mekanisme permintaan dan penawaran [1].

Sebuah negara yang terus mengimpor minyak perlu mengatur strategi dalam pemakaian dan pendapatan minyak, sehingga negara tidak mengalami krisis. Negara diharapkan bisa melihat kemungkinan

kapan kenaikan dan penurunan harga minyak terjadi. Dengan cara meramalkan harga minyak ini, strategi yang diatur diharapkan dapat memberikan manfaat yang maksimal dan menghindarkan negara dari kemungkinan krisis yang bisa saja terjadi.

Harga minyak dapat berubah setiap hari, maka dari itu dalam penelitian ini menggunakan Metode Deep Learning LSTM (Long Short Term Memory) untuk Time Series Forecasting atau Peramalan Seri Waktu dari perubahan harga minyak. Metode LSTM digunakan karena arsitektur LSTM dapat beradaptasi dengan belajar non-linear dan data seri waktu yang kompleks [2].

Penelitian tentang prediksi harga minyak ini telah dilakukan beberapa kali, diantaranya yaitu, menggunakan metode gabungan antara Artificial Neural Network dan Genetic Algorithm [3], metode gabungan antara Multilayer Perception dan Elman Recurrent Neural Network [4], dan dengan metode Support Vector Machine [5].

Tujuan dari penelitian ini adalah untuk membuat sebuah model LSTM untuk meramalkan harga minyak dunia pada periode waktu tertentu. Model LSTM yang

dibuat terbukti dapat memberikan nilai eror yang kecil sehingga dianggap mampu meramalkan harga minyak dengan baik.

2. BAHAN DAN METODE

2.1. Dataset Penelitian

Dataset yang digunakan pada penelitian ini adalah dataset dari Brent Oil Prices [6]. Dataset ini berisi seri waktu harga minyak Brent dari 17 Mei 1987 sampai 25 Februari 2020 berjumlah 8321 *record* yang akan terus di perbarui setiap minggu. Pada dataset ini terdapat 2 atribut. Atribut pertama adalah *Date* dan kedua adalah *Price*. Dataset bisa dilihat pada Tabel 1.

Tabel 1. Dataset harga minyak Brent

No	Date	Price
1	05/20/1987	18.63
2	05/21/1987	18.45
3	05/22/1987	18.55
4	05/25/1987	18.60
5	05/26/1987	18.63
...
8317	02/19/2020	59.72
8318	02/20/2020	59.72
8319	02/21/2020	58.60
8320	02/24/2020	56.04
8321	02/25/2020	56.71

2.2. Preprocessing

Pada tahap *preprocessing* dibagi menjadi beberapa tahap. Tahap yang pertama adalah mengubah tipe data pada kolom *Date* atau tanggal yang sebelumnya memiliki tipe data objek menjadi tipe data *datetime64*. Perubahan tipe data ini dilakukan untuk memudahkan program membaca tipe data pada dataset. Setelah itu atribut *date* diubah menjadi indeks. Perubahan atribut *date* menjadi indeks adalah untuk memberi label pada setiap baris data. Setelah dataset dilakukan *preprocessing* tahap pertama hasilnya bisa dilihat pada Tabel 2.

Tabel 2. Data harga minyak Brent setelah *preprocessing* pertama

Date	Price
1987-05-20	18.63
1987-05-21	18.45
1987-05-22	18.55
1987-05-25	18.60
1987-05-26	18.63
...	...
2020-02-19	59.72
2020-02-20	59.72
2020-02-21	58.60
2020-02-24	56.04
2020-02-25	56.71

Tahap *preprocessing* kedua adalah normalisasi data. Normalisasi data adalah salah satu teknik *preprocessing* yang digunakan untuk menangani atribut yang diluar batas [7]. Normalisasi pada penelitian ini dilakukan dengan dua metode pada model yang berbeda untuk mendapatkan model dengan hasil peramalan yang bagus.

Metode Normalisasi pertama adalah dengan menggunakan *MinMaxScaler* dengan skala 0 sd. 1. *MinMaxScaler* disini merubah data yang diinginkan menjadi di antara rentang 0 sampai dengan 1.

Metode Normalisasi kedua adalah dengan menggunakan *StandardScaler*[8]. *StandardScaler* mentransformasikan dataset sehingga nilai rata-rata distribusi yang dihasilkan adalah nol dan standar deviasi adalah satu. Nilai yang diubah diperoleh dengan mengurangi nilai rata-rata dari nilai asli dan membaginya dengan standar deviasi.

Tahap *preprocessing* ketiga adalah membagi dataset menjadi data uji dan data latih. Pada penelitian ini untuk memperoleh hasil terbaik data uji dan data latih dibagi menjadi dua pembagian. Pembagian pertama adalah 70% data latih dan 30% data uji, dan pembagian kedua adalah 80% data latih dan 20% data uji.

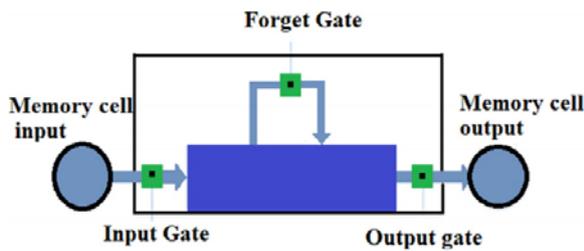
2.3. Long Short Term Memory (LSTM)

LSTM adalah salah satu jenis pemrosesan lain untuk RNN yang diciptakan oleh Hochreiter & Schmidhuber pada tahun 1997 [9]. Metode ini cocok untuk data sekuens. Metode ini mencoba memberikan prediksi menggunakan langkah demi langkah data urutan. Keunggulan LSTM cocok untuk prediksi deret waktu dan sangat bagus dalam menangani data nonlinier yang sangat besar [10].

LSTM berisi unit khusus yang disebut blok memori di lapisan tersembunyi yang berulang. Blok memori berisi sel-sel memori dengan hubungan yang menyimpan keadaan jaringan sementara di samping unit-unit penggandaan khusus yang disebut gerbang pengontrol aliran informasi [11]. Setiap blok memori dalam arsitektur LSTM berisi tiga jenis gerbang yaitu:

1. *Input Gate*: mengontrol aliran aktivasi *input* ke dalam sel memori.
2. *Output Gate*: mengontrol aliran *output* dari aktivasi sel ke seluruh jaringan.
3. *Forget Gate*: menskala keadaan didalam sel sebelum menambahkannya sebagai *input* ke sel melalui koneksi sel yang berulang sendiri, oleh karena itu secara adaptif melupakan atau mengatur ulang memori sel [11].

Pada gambar 1 [12] Gerbang LSTM membantu menangkap dependensi jangka panjang dan jangka pendek dari *input* data seri waktu dan mencegah penurunan gradien atau meledaknya transmisi informasi. Kunci LSTM untuk mewujudkan memori jangka panjang terletak pada menjaga informasi *input* dari setiap langkah dalam unit memori. Status lapisan tersembunyi dari setiap *output* berisi semua informasi *input* sebelum keadaan sekarang [13].



Gambar 1. Arsitektur LSTM

Jaringan LSTM seharusnya bisa mempelajari fitur data latih jangka panjang dan jangka pendek. Untuk mencapai tujuan ini, tipe data *input* memiliki relevansi dengan efektivitas pembelajaran. Jika data yang disediakan mengarah ke arah yang salah atau tidak cukup untuk membuat fitur menjadi jelas, maka LSTM akan belajar sesuai dengan itu dan tidak akan memprediksi atau memperkirakan secara akurat [14].

3. HASIL DAN PEMBAHASAN

Tabel 3. Perbandingan parameter

Model No.	Metode Normalisasi	Parameter				
		Perbandingan Data		Total Layer	Batch Size	Total Epoch
		Latih	Uji			
1	<i>MinMaxScaler</i>	70	30	3	15	20
2	<i>MinMaxScaler</i>	70	30	4	15	20
3	<i>MinMaxScaler</i>	80	20	3	32	30
4	<i>MinMaxScaler</i>	80	20	3	64	30
5	<i>StandardScaler</i>	80	20	3	8	30
6	<i>StandardScaler</i>	70	30	3	4	30

Setelah pelatihan selesai file model setiap percobaan disimpan dalam format .h5 untuk nanti dapat dipanggil lagi untuk peramalan.

Pada penelitian ini untuk mengetahui seberapa bagus model yang dibuat menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE) sebagai metode evaluasi percobaan.

RMSE secara matematis, rumusnya ditulis dengan Persamaan 1

Pada penelitian ini menggunakan bahasa pemrograman Python. Python digunakan karena menyediakan banyak pustaka untuk melakukan pemrosesan *Deep Learning* seperti LSTM.

Setelah melakukan *preprocessing* data kemudian proses selanjutnya adalah membangun model untuk peramalan. 6 Model dengan beberapa parameter yang berbeda telah dibangun untuk mengetahui model yang memberikan performa terbaik untuk peramalan harga minyak.

Ada juga parameter yang tidak diubah seperti banyak *units* dalam LSTM, *dropout*, *optimizer* dan *callbacks*. *Optimizer* menggunakan *adam*, *unit* dalam LSTM berjumlah 60, *dropout* sebesar 0.1 dan *callbacks* menggunakan *ReduceLROnPlateau* [15]. Fungsi *ReduceLROnPlateau* akan mengurangi *learning rate* jika *learning rate* pada *epoch* sekarang menyebabkan model tidak menunjukkan perkembangan terhadap apa yang diawasi.

Parameter - parameter yang berbeda pada model dapat dilihat pada Tabel 3.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - q_i)^2} \quad (1)$$

dimana y_i adalah nilai prediksi dan q_i adalah nilai asli.

MAE adalah rata-rata nilai absolut dari jumlah penyimpangan semua pengamatan individu dari rata-rata aritmatika. Rumusnya dituliskan pada Persamaan 2.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - q_i| \quad (2)$$

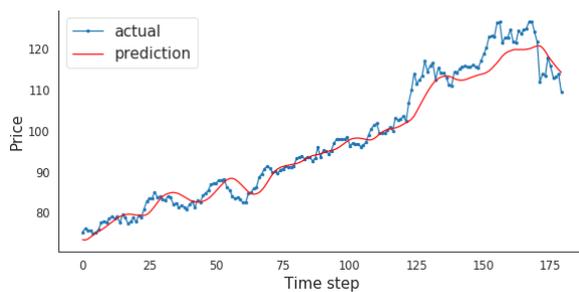
Hasil evaluasi percobaan dapat dilihat pada Tabel 4

Tabel 4. Hasil percobaan model LSTM

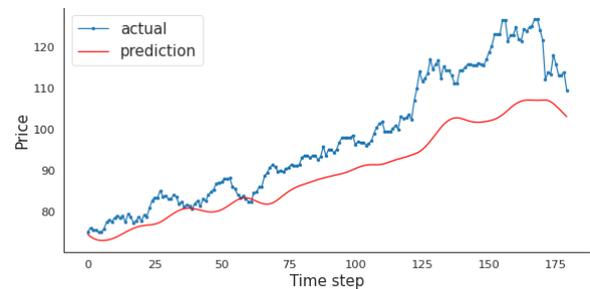
Model No.	Loss	MAE	RMSE
1	0.00046735	1.2727834120285700	1.9235046784378689
2	0.0008713	4.0982606947692295	4.986727484132979
3	0.00051355	1.21415301387135	1.9186450688556458
4	0.000546	1.2910977920685605	2.014103478099587
5	0.0171	0.013022615377289829	0.018599384625417048
6	0.0100	1.5864032508988475	2.253581372417259

Untuk melakukan prediksi dilakukan memuat berkas model.h5. Prediksi harga minyak menggunakan model ini bisa dilakukan dengan kurun waktu tak terbatas, namun untuk mengetahui apakah hasil peramalan bagus atau tidak peramalan dilakukan selama 180 hari kedepan dengan cara membandingkan dengan data asli.

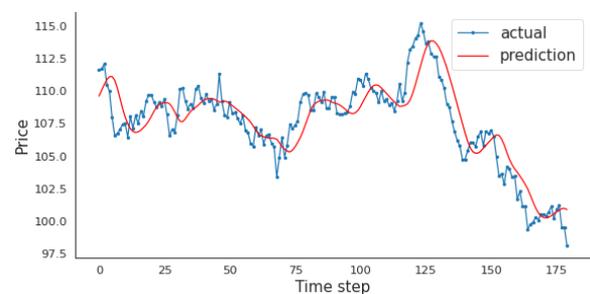
Pada hasil peramalan model 1 pada Gambar 2, Metode Normalisasi yang digunakan adalah *MinMaxScaler*, pembagian data uji dan data latih sebesar 70% : 30 %, total *layer* sebanyak 3, *batch size* sebesar 15 dan *epoch* 20 menunjukkan secara keseluruhan model sudah bisa meramalkan harga dengan bagus namun pada beberapa titik kenaikan dan penurunan mengalami keterlambatan. Perbedaan juga terdapat pada setiap puncak harga pada waktu tertentu.

**Gambar 2.** Peramalan Model 1

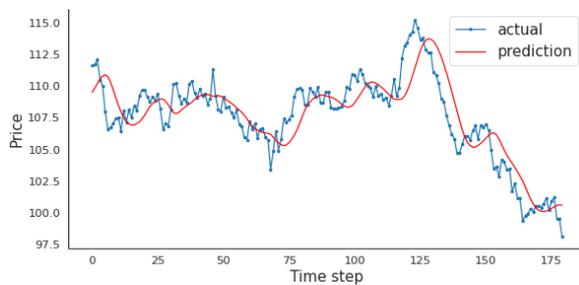
Pada hasil peramalan model 2 pada Gambar 3 semua parameter mirip dengan model 1 namun yang membedakan ada penambahan layer. Total *layer* pada model 2 sebanyak 4. Hasil peramalan model 2 menunjukkan ketidaksesuaian prediksi dengan data yang sebenarnya.

**Gambar 3.** Peramalan Model 2

Pada model 3 yang ditunjukkan pada Gambar 4 dengan metode normalisasi *MinMaxScaler*, perbandingan data latih dan data uji sebesar 80% : 20 %, total *layer* sebanyak 3, *batch size* 32, dan *epoch* 30 menunjukkan hasil yang kurang lebih sama dengan model 1 namun lebih baik. Hasil peramalan model 3 secara keseluruhan sudah bisa meramalkan harga minyak dengan baik namun ada sedikit keterlambatan pada saat harga mengalami penurunan atau kenaikan. Dan beberapa puncak harga pada waktu tertentu juga memiliki perbedaan.

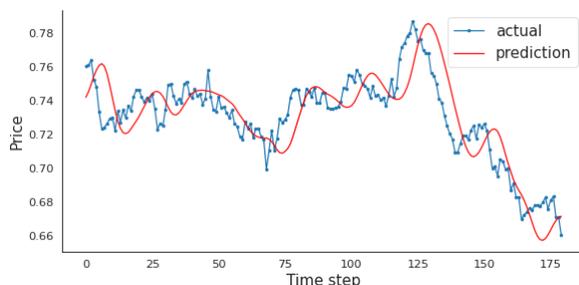
**Gambar 4.** Peramalan Model 3

Pada hasil peramalan model 4 pada Gambar 5 perbedaannya hanya besaran *batch size* menjadi 64 dari model ke 3 yang besarnya 32. Hasil peramalan model 4 ini mirip dengan model 3. Yang menjadi pembedanya adalah hasil evaluasi *Loss*, *MAE* dan *RMSE*.



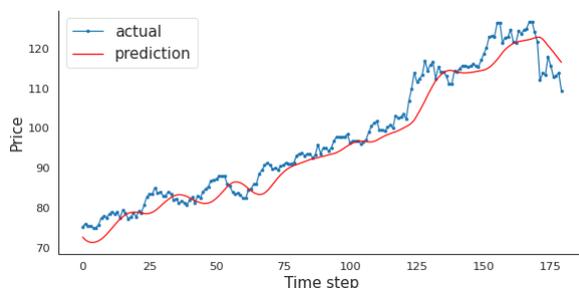
Gambar 5. Peramalan Model 4

Pada hasil peramalan model 5 pada Gambar 6 parameter yang dibedakan dari model 4 adalah metode normalisasi dan besaran *batch size*. Pada model 5 ini menggunakan metode normalisasi *StandardScaler* dan *batch size* sebesar 8. Hasil prediksi menunjukkan peramalan yang sudah cukup baik, meskipun masih ada keterlambatan kenaikan dan penurunan harga. Harga minyak puncak pada waktu tertentu juga memiliki nilai yang hampir sama antara data asli dengan data prediksi. Dan model 5 ini menghasilkan MAE dan RMSE terkecil dari semua model.



Gambar 6. Peramalan Model 5

Pada hasil peramalan model 6 pada Gambar 7 menggunakan metode normalisasi *StandardScaler*, perbandingan data latih dan data uji sebesar 70% : 30 %, total *layer* sebanyak 3, *batch size* sebesar 4 dan *epoch* 30. Hasil peramalan model 6 kurang lebih sama dengan model 1 namun tidak lebih baik.



Gambar 7. Peramalan Model 6

4. KESIMPULAN

Beberapa model LSTM yang dibangun terbukti dapat meramalkan harga minyak Brent dengan baik. Model yang memiliki hasil terbaik adalah model 5 dengan MAE dan RMSE terkecil. Perbedaan gambar

peramalan disebabkan karena perbedaan perbandingan data uji dan data latih. Penambahan *layer* pada model LSTM ke 2 tidak menyebabkan hasil peramalan menjadi baik namun menjadikan model memiliki eror yang paling besar dari model yang lain. Ini bisa disebabkan karena faktor parameter yang kurang sesuai. Perbedaan *epoch* dan *batch size* pada penelitian ini tidak berdampak besar pada pengurangan eror. Namun perbedaan metode normalisasi menjadikan pengurangan yang lumayan besar pada perubahan eror dan menambah *loss* menjadi sedikit lebih besar.

5. DAFTAR PUSTAKA

- [1] P. Ponadi, D. Amboningtyas, and A. Fathoni, "ANALYSIS OF THE EFFECT OF WORLD OIL PRICE INCREASE, AMOUNT OF CIRCULAR MONEY, AND EXCHANGE ON INFLATION IN INDONESIA (Case Study of Mining Companies in the 2013-2017 Period)," *J. Manage.*, vol. 5, no. 5, 2019.
- [2] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, 2019.
- [3] M. R. Mahdiani and E. Khamehchi, "A modified neural network model for predicting the crude oil price," *Intellect. Econ.*, vol. 10, no. 2, pp. 71–77, 2016, doi: 10.1016/j.intele.2017.02.001.
- [4] J. Wang and J. Wang, "Forecasting energy market indices with recurrent neural networks: Case study of crude oil price fluctuations," *Energy*, vol. 102, pp. 365–374, 2016.
- [5] L. Yu, X. Zhang, and S. Wang, "Assessing potentiality of support vector machine method in crude oil price forecasting," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 13, no. 12, pp. 7893–7904, 2017, doi: 10.12973/ejmste/77926.
- [6] M. Abusalah, "Brent Oil Price," 2020. <https://www.kaggle.com/mabusalah/brent-oil-prices> (accessed Mar. 20, 2020).
- [7] S. Jain, S. Shukla, and R. Wadhvani, "Dynamic selection of normalization techniques using data complexity measures," *Expert Syst. Appl.*, vol. 106, pp. 252–262, 2018.
- [8] Scikit-learn.org, "6.3. Preprocessing Data — Scikit-Learn 0.23.1 Documentation," 2020. <https://scikit-learn.org/stable/modules/preprocessing.html> (accessed Mar. 23, 2020).
- [9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [10] Y. Sudriani, I. Ridwansyah, and H. A. Rustini, "Long short term memory (LSTM) recurrent neural network (RNN) for discharge level prediction and forecast in Cimandiri river, Indonesia," in *IOP Conference Series: Earth and Environmental Science*, 2019, vol. 299, no.

- 1, p. 12037.
- [11] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting," *Procedia Comput. Sci.*, vol. 135, pp. 89–98, 2018.
 - [12] V. K. R. Chimmula and L. Zhang, "Time Series Forecasting of COVID-19 transmission in Canada Using LSTM Networks," *Chaos, Solitons & Fractals*, p. 109864, 2020.
 - [13] Y. Ding, Y. Zhu, J. Feng, P. Zhang, and Z. Cheng, "Interpretable Spatio-Temporal Attention LSTM Model for Flood Forecasting," *Neurocomputing*, 2020.
 - [14] S. Muzaffar and A. Afshari, "Short-term load forecasts using LSTM networks," *Energy Procedia*, vol. 158, pp. 2922–2927, 2019.
 - [15] M. Zaheer, S. Reddi, D. Sachan, S. Kale, and S. Kumar, "Adaptive methods for nonconvex optimization," in *Advances in neural information processing systems*, 2018, pp. 9793–9803.