

Prediksi Arah Harga Bitcoin Berdasarkan Manipulasi Metode *Long Short-Term Memory* (LSTM)

Asep Saepulrohman^{1*}, Shandy Zolla Pratama²

^{1,2}Program Studi Ilmu Komputer, Universitas Pakuan

*e-mail: asepspl@unpak.ac.id

Diterima: 14 Februari 2023, disetujui: 27 Februari 2023, dipublikasi: 12 Maret 2023

Abstract: Bitcoin is one of the new currencies in online transactions that was found to be one of the cryptocurrencies that has fluctuating volatility, so researchers are looking for the determinants of bitcoin price. In this paper, we adopt a Long Short-Term Memory Networks (LSTM) approach to evaluate the effect of investor attention on bitcoin returns by constructing an aggregate investor attention proxy. In addition, LSTM is a method of Recurrent Neural Network (RNN) to deal with the problem of long-term dependence on RNN. The main objective of this research is to design, implement, and visualize Bitcoin Price Prediction using the LSTM method. The data used in this study were daily bitcoin cover data and tweets taken from twitter. The measure of the accuracy of the model used is the Root Mean Square Error (RMSE). The results show that the RMSE value was quite small, meaning that it LSTM was quite good Bitcoin prediction model.

Keywords: Bitcoin; LSTM; RMSE; RNN; cryptocurrency

Abstrak: Bitcoin merupakan salah mata uang baru dalam transaksi online yang ditemukan sebagai salah satu cryptocurrency yang memiliki volatilitas yang berubah-ubah, sehingga para peneliti mencari faktor penentu harga bitcoin. Dalam makalah ini, Penelitian ini menggunakan pendekatan Long Short-Term Memory Networks (LSTM) untuk mengevaluasi pengaruh perhatian investor terhadap pengembalian bitcoin dengan membangun proxy perhatian investor agregat. Selain itu juga, LSTM adalah salah satu metode Recurrent Neural Network (RNN) untuk menangani masalah ketergantungan jangka panjang pada RNN. Tujuan utama penelitian ini adalah merancang, mengimplementasikan, dan memvisualisasikan Prediksi Harga Bitcoin menggunakan metode LSTM. Data penelitian ini yaitu data penutup bitcoin harian dan tweet yang diambil dari twitter [xxx]. Ukuran keakuratan model yang digunakan adalah Root Mean Square Error (RMSE). Hasil penelitian menunjukkan bahwa nilai RMSE cukup kecil artinya cukup baik dalam memodelkan prediksi bitcoin.

Kata Kunci: Bitcoin; LSTM; RMSE; RNN; cryptocurrency

1. PENDAHULUAN

Diketahui bahwa pasar keuangan memiliki karakteristik yang kompleks diakibatkan adanya unsur atau elemen yang saling berinteraksi di pasar modal dan fitur nonlinear [1]. Banyak penelitian sebelumnya [2]-[5] yang menganalisis tentang karakteristik pasar dengan data deret waktu keuangan, diantaranya adalah memprediksi data keuangan yang sangat dipengaruhi oleh faktor ekonomi, politik, informasi asimetris dan faktor alam. Saat ini, *bitcoin* muncul di pasar digital sangat cepat dimana *bitcoin* merupakan mata uang elektronik atau *cryptocurrency* yang dibuat pada tahun 2009 oleh Satoshi Nakamoto karena merupakan bentuk rasa kekecewaan pada krisis ekonomi tahun 2008.

Bitcoin banyak digunakan untuk transaksi *online* tanpa menggunakan perantara dengan sistem kerja *bitcoin* yaitu menggunakan sistem *peer-to-peer* tanpa penyimpanan terpusat atau administratornya tunggal dimana Departemen Keuangan Amerika Serikat menyebut *bitcoin* sebagai sebuah mata uang yang terdesentralisasi. Pada awal Januari 2013, *bitcoin* ini hanya dihargai perkoinnya adalah US \$13 per koin dalam 1 BTC [6]-[9]. Harga tersebut kemudian melonjak tinggi menjadi lebih dari \$1.100 per koin pada bulan Desember di tahun yang sama, tetapi tidak lama kemudian harga *bitcoin* tersebut turun menjadi setengahnya dalam beberapa jam. Hal ini dikarenakan adanya larangan perdagangan *bitcoin* di China pada waktu itu. Harga fluktuatif yang terjadi pada *bitcoin* ini akan menjadi masalah besar apabila seseorang ingin melakukan perdagangan *bitcoin*, oleh karena itu pada era ini sudah banyak aplikasi – aplikasi maupun *website trading* yang dapat membantu pedagang tersebut dalam melakukan perdagangan *bitcoin*, salah satunya adalah Investing.com.

Ekplorasi data nonlinear digunakan sebagai dasar untuk memprediksi data *time series* untuk membentuk model sehingga data tersebut bisa dijadikan acuan untuk memprediksi nilai berikutnya. Jenis data prediksi *time series* lebih jika dibandingkan pemodelan prediktif regresi karena prediksi *time series* menambah kompleksitas ketergantungan urutan antar *variable*. Dalam penelitian ini ditunjukkan bahwa *Recurrent Neural Network* (RNN) [10] telah terbukti berhasil digunakan untuk prediksi data *time series* karena RNN mampu menggunakan informasi yang telah didapatkan sebelumnya yang urutannya panjang atau *sequence*-nya beragam. Pembangunan sistem ini dibuat dengan metode *Recurrent Neural Network* dengan menggunakan variasi dari RNN yaitu *Long Short-Term Memory* (LSTM) [11]. Ada keunggulan LSTM jika dibandingkan dengan RNN, yaitu adanya arsitektur yang menjadikan *output* yang nantinya akan diproses kembali menjadi *input*. Untuk membuat prediksi arah harga *bitcoin* harus diperlukan data historis terlebih dahulu untuk dapat menunjukkan pola peristiwa masa lalu. Data tersebut digunakan data harga *bitcoin* per hari yang diperoleh dari portal *investing.com*.

2. METODOLOGI PENELITIAN

Dalam penelitian ini metode yang digunakan adalah menggunakan *database* atau disebut juga *Knowledge Discovery in Database* (KDD) dimana periode waktunya dari bulan Januari 2008 sampai dengan bulan September 2021. KDD merupakan metode untuk memperoleh pengetahuan dari database yang ada [12]-[15] dengan harga yang digunakan adalah harga *bitcoin*. Tahapan KDD adalah sebagai berikut:

Data Selection (DS)

DS merupakan tahapan awal untuk memfilter data, dimana sebelum memulai tahap menemukan informasi dalam KDD, data harus diseleksi adalah data yang diperoleh sebelumnya. Data yang lulus seleksi akan digunakan untuk proses *data mining*, kemudian data tersebut disimpan ke dalam suatu *file* yang terpisah dari *database*.

Pre-processing (PP)

Tahapan ini adalah proses *data mining*, ada hal yang perlu dilakukan terlebih dahulu *pre-processing* pada data inti KDD. Proses *pre-processing* meliputi menghapus data duplikat, memeriksa data yang tidak konsisten dan memperbaiki kesalahan pada data. Data yang sudah di *pre-processing* juga dilakukan proses *enrichment*, yaitu menambahkan data eksternal atau informasi penting lain yang relevan dan diperlukan untuk KDD.

Transformation (Tr)

Perubahan data atau transformasi data yang dipilih dari data tersebut cocok untuk proses *data mining*. Proses transformasi dalam KDD merupakan proses yang sangat penting dan bergantung pada jenis atau model informasi yang akan diambil dalam *database*.

Data Mining

Selanjutnya adalah *data mining*, dalam tahapan ini adalah proses untuk menemukan pola data yang dipilih dengan menggunakan teknik atau metode yang tepat. Dengan demikian, pemilihan algoritma atau metode yang tepat sangat bergantung pada tujuan dan seluruh proses KDD.

Evaluation (Ev)

Terakhir adalah melakukan evaluasi dimana dalam tahapan ini data atau informasi yang diperoleh dari hasil data mining harus menampilkan informasi yang mudah dipahami oleh lain. Langkah ini merupakan bagian dari proses KDD yang disebut *interpretation*. Langkah ini termasuk memeriksa apakah ada model atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada. LSTM adalah salah satu metode dari RNN yang menangani masalah ketergantungan dalam waktu jangka panjang pada RNN. Dalam hal ini, RNN perulangan model hanya menggunakan satu *layer* tunggal yang sederhana, yaitu *layer tanh*. Dimana LSTM untuk menangani masalah ketergantungan memiliki dua *cell* yaitu *cell state* dan *hidden state*. *Cell state* merupakan jaringan utama yang memungkinkan data mengalir pada dasarnya. Data dapat dihapus melalui *sigmoid gates*. LSTM dirancang untuk menghindari masalah ketergantungan jangka panjang [15]-[19]. LSTM memiliki 4 persamaan dalam layer-nya, berikut rumusnya:

1. Forget Gate

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \quad (1)$$

2. Input Gate

$$I_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (2)$$

3. Memory Update

$$\tilde{c}_t = \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (3)$$

$$c_t = f_t * c_t + i_t * \tilde{c}_t \quad (4)$$

4. Output Gate

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (5)$$

$$\tilde{h}_t = o_t * \tanh(c_t) \quad (6)$$

Keterangan:

- f_t : Lupakan gerbang (*Forget Gate*)
- I_t : *Input* gerbang
- c_t : *Update Cell State*
- o_t : *Output* gerbang
- W : *Weight*
- b : *Bias*
- σ : *Sigmoid*
- h_t : Nilai *Output*
- x_t : Nilai *Input*
- h_{t-1} : Nilai *Output Cell* Sebelumnya
- C_{t-1} : *Cell State* Sebelumnya

Fungsi aktivasi yang digunakan secara umum merupakan *sigmoid*, *tanh* dan *softmax*, tetapi yang digunakan untuk penelitian kali ini hanya menggunakan *sigmoid* dan *tanh* [1]. Penjelasan mengenai fungsi aktivasi tersebut adalah sebagai berikut:

1. *Sigmoid*

Fungsi aktivasi yang pertama adalah *sigmoid*, fungsi tersebut merupakan fungsi yang *nonlinear*. Dimana input untuk fungsi aktivasi adalah berupa bilangan riil, sedangkan *output* dari fungsi tersebut memiliki nilai *range* ada pada interval 0 dan 1. Fungsi aktivasi *sigmoid*

$$sigmoid(x) = \frac{1}{1+e^{-x}} \quad (7)$$

2. *Tanh*

Fungsi aktivasi *tanh* merupakan fungsi *nonlinear*. Masukan untuk fungsi aktivasi ini berupa bilangan *real* dan *output* dari fungsi tersebut memiliki *range* antara -1 sampai 1. Perhitungan fungsi aktivasi *tanh* sebagai berikut:

$$sigmoid(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (8)$$

Metode alternatif dalam mengevaluasi hasil peramalan dengan tingkat akurasi tinggi dapat menggunakan *Root Mean Square Error* (RMSE) yaitu metode alternatif untuk perkiraan suatu model. Nilai yang dihasilkan RMSE adalah akar dari data hasil nilai rata-rata kuadrat dari jumlah kesalahan pada model prediksi [19].

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (y_i - \tilde{y}_i)^2} \quad (9)$$

Keterangan:

- \tilde{y} : Data aktual
- y : Data prediksi

Galat persentase:

$$\varepsilon = \left(\frac{\alpha - a}{a} \right) \times 100\% \quad (10)$$

Keterangan:

- α : Nilai prediksi
- a : Nilai aktual

3. HASIL DAN PEMBAHASAN

Hasil dari penelitian yang dibangun menggunakan jaringan LSTM dengan 30 *hidden unit* yang ada pada *layer* dan 60 kali *training*. Dengan menggunakan 2 hal tersebut jaringan LSTM ini memaksa jaringan untuk mencari *pattern* dengan tujuan memprediksi harga di masa mendatang. Beberapa hasil dari prediksi harga dan sentimen pada Gambar 1.



Gambar 1. Hasil prediksi pertama

Data yang diperoleh dari Januari 2018 sampai dengan Juli 2021 dengan data penelitian tersebut dari data train, data tes, dan prediksi dan terjadi kenaikan signifikan terjadi pada bulan Januari 2021. Kemudian data untuk memprediksi data dilakukan seperti pada Gambar 2.



Gambar 2. Hasil prediksi kedua

Hasil prediksi kedua masih tetap dari Januari 2018 sampai dengan bulan Juli 2021 dengan data tersebut dari data train, data tes, dan hasil prediksi dan terjadi kenaikan signifikan masih terjadi pada bulan Januari 2021. Kemudian data untuk memprediksi data dilakukan seperti pada Gambar 3.



Gambar 3. Hasil prediksi ketiga

Data hasil prediksi ketiga yang diperoleh dari bulan Januari 2021 sampai dengan Juli 2021 dengan data tersebut dari data train, data tes, dan prediksi dan terjadi perbedaan dengan hasil pada data hasil prediksi kedua dimana kenaikan terjadi pada bulan Januari 2021, tetapi terjadi penurunan pada bulan yang sama. Kemudian dilakukan prediksi keempat dilakukan seperti pada Gambar 4.



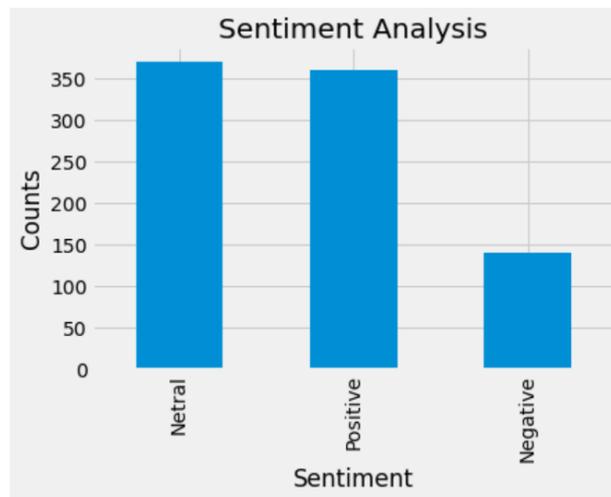
Gambar 4. Hasil prediksi keempat

Data yang diperoleh dari Januari 2021 - Juli 2021 memiliki kemiripan dengan data yang diperoleh pada prediksi keempat. Kemudian dilakukan prediksi tahap lima seperti pada Gambar 5.



Gambar 5. Hasil prediksi kelima

Data yang diperoleh dari bulan Januari 2002 sampai dengan bulan Juli 2021 dengan data tersebut dari data train, data tes, dan prediksi dan terjadi kenaikan signifikan terjadi pada tahun 2021 dari bulan Januari 2021 dan terjadi penurunan kembali yang di bulan Juli. Kemudian dari beberapa kali prediksi dapat dianalisis dimana sentimen analisisnya pada Gambar 6.



Gambar 6. Hasil sentimen analisis

Dikarenakan pada model ini terdapat 3 diantaranya menunjukkan tren positif, dan hasil sentimen pasar pun menunjukkan tren netral ke positif, maka kita dapat menyimpulkan bahwa harga bitcoin masih dapat naik kembali setelah penurunan ini. Kemudian model ini selanjutnya dihitung *Root Mean Square Error* (RMSE). Tabel hasil RMSE ada pada Tabel 1.

Tabel 1. Hasil RMSE

Hasil Uji ke-	RMSE
1	1205,850
2	1311,820
3	1826,129
4	1112,641
5	1641,869
Rata-rata	1419,6618

Berdasarkan Tabel 1, terlihat hasil uji ke-3 memiliki hasil terburuk dengan 1826.129 untuk RMSEnya, dan yang terbaik ada pada hasil uji ke-4 dimana memiliki hasil 1112.641 untuk RMSEnya. Rata-rata dari ke 5 uji hasil RMSEnya adalah 1419.6618. Galat presentasenya dapat dilihat sebagai berikut:

$$\varepsilon = \left(\frac{\alpha - a}{a} \right) \times 100\%$$

$$\varepsilon = \left(\frac{47637.6618 - 46218}{46218} \right) \times 100\%$$

$$\varepsilon = 3.07\%$$

Dikarenakan sebuah prediksi harga ini dapat menjadi salah satu pertimbangan jika ingin berinvestasi di *bitcoin* dengan melihat prediksi harga yang didapat menggunakan data harga *time-series* nya, kita dapat memprediksi apakah bagus untuk mulai berinvestasi sekarang. Metode LSTM dapat memprediksi harga ke depan, berdasarkan analisis hasil prediksi yang didapat harga masih menunjukkan tren positif. Bulan Oktober 2021 sampai Desember 2021 yang menunjukkan grafik harga meningkat, maka dapat disimpulkan prediksi harga *bitcoin* ini masih aman untuk 3 bulan ke depan tetapi masih harus diperhatikan ada siklus penurunan setiap akhir tahun. Model ini dapat digunakan untuk 1 tahun kedepan, dengan asumsi tidak ada kelonjakan harga yang signifikan.

4. KESIMPULAN

Dalam penelitian data yang digunakan adalah sebanyak 1369 data yang diambil dari *investing.com* pada bulan Oktober 2021 dan untuk akurasi yang digunakan di penelitian ini menggunakan *Root Mean Square Error* (RMSE). Hasil prediksi harga didapatkan akurasi RMSE rata – rata 1419,6618 serta galat persentasenya 3,07%. Data sentimen pasar yang diambil dari Twitter pada bulan Desember 2021, didapatkan data sebanyak 872 data. Hasil sentimen analisisnya mendapatkan 361 positive tweets, 370 netral tweets, dan 141 *negative tweets*.

DAFTAR PUSTAKA

- [1] Koo, E., & Kim, G. (2021). Prediction of Bitcoin price based on manipulating distribution strategy. *Applied Soft Computing*, 110, 107738.
- [2] Mitsuru Iwamura, Yukinobu Kitamura, Tsutomu Matsumoto, Kenji Saito, Can we stabilize the price of a cryptocurrency?: Understanding the design of Bitcoin and its potential to compete with central bank money, *Hitotsubashi J. Econ.* (2019) 41–60.
- [3] David Yermack, Is Bitcoin a real currency? An economic appraisal, in: *Handbook of Digital Currency*, Elsevier, 2015, pp. 31–43.
- [4] Dirk G. Baur, Kihoon Hong, Adrian D. Lee, Bitcoin: Medium of exchange or speculative assets? *J. Int. Financ. Mark. Inst. Money* 54 (2018) 177–189.
- [5] Vladimir Stojanovic, Dragan Prsic, Robust identification for fault detection in the presence of non-Gaussian noises: application to hydraulic servo drives, *Nonlinear Dynam.* 100 (2020) 2299–2313.
- [6] S. Wahyudi and C. Imron, “Super Edge-Magic Pada[1] Faadilah, A. 2020. Analisis Sentimen Pada Ulasan Aplikasi Tokopedia di Google Play Store Menggunakan Metode Long Short Term Memory. Skripsi. Jakarta: UIN Syarif Hidayatullah Jakarta.
- [7] Fariska, P., Nugraha, & Rohandi, M. M. A. 2020. Hubungan Sentimen Investor, Volume Perdagangan dan Kebijakan Moneter Pada Perkembangan Pasar Modal di Indonesia. *Jurnal Manajemen dan Bisnis*, 17(1) e-ISSN 2599-0039.
- [8] Fatah, H., Subekti, A., 2018. Prediksi Harga Cryptocurrency Dengan Metode K-Nearest Neighbours. *Jurnal PILAR Nusa Mandiri* Vol. 14, No. 2. P-ISSN: 1978-1946, E-ISSN: 2527-6514.
- [9] Ilmawan, D., Warsito, B., & Sugito, S. 2020. Penerapan Artificial Neural Network Dengan Optimasi Modified Artificial Bee Colony Untuk Meramalkan Harga Bitcoin Terhadap Rupiah. *Jurnal Gaussian*, 9(2), 135-142.
- [10] Juanda, R. A., Jondri, J., & Rohmawati, A. A. 2018. Prediksi Harga Bitcoin Dengan Menggunakan Recurrent Neural Network. *eProceedings of Engineering*, 5(2).
- [11] Le, X. H., Ho, H. V., Lee, G., & Jung, S. 2019. Application of long short-term memory (LSTM) neural network for flood forecasting. *Water*, 11(7), 1387.
- [12] Mardi, Y. 2017. Data Mining: Klasifikasi Menggunakan Algoritma C4.5. *Edik Informatika*, 2(2), 213-219.
- [13] Mehta, P., Sasikala, E. 2020. Prediction of Bitcoin using Recurrent Neural Network. *International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878 (Online), Volume-8 Issue-6.
- [14] Nurdiansyah, A., Furqon, M. T., & Rahayudi, B. 2019. Prediksi Harga Bitcoin Menggunakan Metode Extreme Learning Machine (ELM) dengan Optimasi Artificial Bee Colony (ABC). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* e-ISSN, 2548, 964X. [10] A. Rorres, *An Introduction to Linear Algebra*, 4 Edition. London: Pearson, 2017.
- [15] Rohmawati, F., Rohman, M. G., & Mujilahwati, S. 2017. Sistem Prediksi Jumlah Pengunjung Wisata Wego Kec. Sugio Kab. Lamongan Menggunakan Metode Fuzzy Time Series. *Joutica*, 2(2).

- [16] Salwa, N., Tatsara, N., Amalia, R., & Zohra, A. F. 2018. Peramalan Harga Bitcoin Menggunakan Metode ARIMA (Autoregressive Integrated Moving Average). *Journal of Data Analysis*, 1(1), 21-31.
- [17] Turban, E., Delen, D., & Sharda, R. 2018. *Business Intelligence, Analytics, and Data Science: A Managerial Perspective*. Harlow ; Munich: Pearson Prentice Hall.
- [18] Winata, W. 2018. *Prakira Suhu Udara Rata-Rata Kota-Kota Besar Dunia Menggunakan Metode Long Short-Term Memory*. Skripsi. Medan: Universitas Sumatera Utara.
- [19] Wiranda, L., & Sadikin, M. 2020. Penerapan Long Short Term Memory Pada Data Time Series Untuk Memprediksi Penjualan Produk Pt. Metiska Farma. *Jurnal Nasional Pendidikan Teknik Informatika: JANAPATI*, 8(3), 184-196.