

**Jurnal Politeknik Caltex Riau**Terbit Online pada laman <https://jurnal.pcr.ac.id/index.php/jkt/>

| e- ISSN : 2460-5255 (Online) | p- ISSN : 2443-4159 (Print) |

Prediksi Harga Bitcoin Menggunakan Metode Random Forest (Studi Kasus: Data Acak Pada Awal Masa Pandemic Covid-19)

Siti Saadah ¹, Haifa Salsabila ²¹Informatics Faculty, Telkom University, email: sitisaadah@telkomuniversity.ac.id²Informatics Faculty, Telkom University, email: haifasalsabila@student.telkomuniversity.ac.id

[1] Abstrak

Bitcoin adalah mata uang digital yang merepresentasikan perubahan penting yang dapat memberikan dampak untuk sektor keuangan. Beberapa tahun belakangan dan terutama pada masa awal pandemic covid-19, pemilihan investasi menggunakan bitcoin sebagai salah satu cryptocurrency semakin banyak digunakan oleh masyarakat di dunia. Bangsa Indonesia sebagai negara berkembang juga memerlukan beberapa penelitian terkait hal ini, guna mempersiapkan terhadap dampaknya sedini mungkin. Oleh sebab itu, pada penelitian ini di bangun sistem prediksi harga bitcoin menggunakan salah satu algoritma pembelajaran mesin. untuk mengetahui pergerakan fluktuasi harga bitcoin. Mengacu kepada karakteristik data bitcoin yang memiliki kecenderungan volatile, maka digunakan metode Random Forest Regression untuk memprediksi harga bitcoin dengan terlebih dahulu dilakukan feature selection terhadap atribut dataset bitcoin. Dengan menggunakan pemodelan Random Forest Regression, diperoleh nilai MAPE sebesar 1.50% dengan akurasi 98.50%. Berdasarkan hasil tersebut, diketahui bahwa algoritma random forest ini merupakan salah satu pemodelan yang dapat menghasilkan performansi yang baik dalam hal prediksi terutama untuk data yang bersifat acak.

Kata kunci: *Prediksi, Harga Bitcoin, Random Forest Regression, MAPE*

[2] Abstract

Bitcoin is a digital currency that represent a significant effect to financial sector. In a recent year and especially in the beginning of covid-19 widespread trend of investor or business using cryptocurrency as a digital transaction emerge immensely from people around the world. Indonesia as a developing country also need a study further to mitigate the turmoil effect of it. Based on that reason, in this research had been conducted prediction system of bitcoin price using one of machine learning algorithm. Regarding into characteristic of bitcoin data that volatile made random forest chosen as the method approach, while there is a feature selection process to do attribute feature selection at the commence. Using model learning from random forest regression derived MAPE at around 1.50% with 98.50% of accuracy. This result indicate that random forest succeeded in producing good performance mainly to predict data random.

Keywords: *Prediction, Bitcoin Price, Random Forest Regression, MAPE*

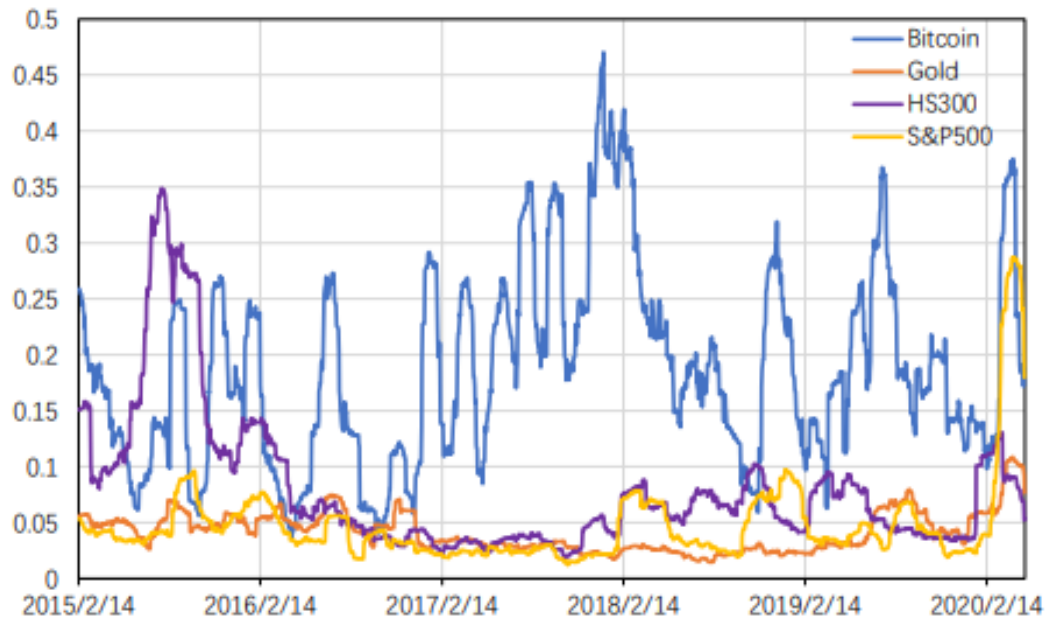
1. Pendahuluan

Mata uang digital yaitu *cryptocurrencies*, dimana salah satunya yang terkenal dan banyak digunakan adalah bitcoin, merupakan salah satu asset finansial yang memegang peranan penting pada *market capitalization* dan pertukaran mata uang. Nilai bitcoin yang senantiasa mengalami fase ketidakstabilan yang tinggi, karena merupakan produk global yang dipengaruhi oleh peraturan di dunia, dibandingkan data asset keuangan yang lain seperti emas, indeks barang dan komoditas menjadi sebuah permasalahan yang perlu di observasi mendalam. Karena fluktuasi dari bitcoin ini tidak dapat dihindari dapat memberikan efek secara tidak langsung kepada kondisi keuangan di sebuah negara, khususnya Indonesia sebagai negara berkembang. Situasi pandemic yang bermula pada tahun 2019, membuat transaksi jual-beli bertransformasi menggunakan layanan digital. Dikarenakan untuk memutus rantai penyebaran Covid-19, bitcoin merupakan salah satu *cryptocurrency* yang saat ini banyak digunakan dan diminati oleh masyarakat dunia karena dijadikan sebagai alat pembayaran digital maupun investasi. Bitcoin pertama kali dikenalkan oleh Satoshi Nakamoto pada Januari 2009. Mata uang bitcoin dikendalikan oleh sistem perangkat lunak terbuka, sehingga siapa saja dapat mengubah sistem tersebut. Pengguna bitcoin menyimpan bitcoinnya pada dompet digital sama seperti akun virtual bank. Setiap transaksi yang dilakukan disimpan pada teknologi *blockchain*. Berdasarkan data pada laman coinmarketcap.com, bitcoin menduduki urutan nomor 1 sebagai *top cryptocurrency* dengan nilai market cap yang sangat tinggi. Pada aplikasi bitcoin, tidak terdapat institusi keuangan yang mengelola ataupun mengawasi transaksi antar pengguna bitcoin sehingga para pengguna bitcoin tidak dapat mengetahui kebijakan yang berlaku mengenai bitcoin, termasuk parameter yang dapat mempengaruhi pergerakan nilai harga bitcoin. Oleh sebab itu, penelitian mengenai prediksi harga bitcoin banyak dikembangkan.

Beberapa penelitian sebelumnya seperti pada penelitian Edwin Sin, dkk [1] memprediksi harga bitcoin menggunakan neural network untuk memprediksi nilai harga bitcoin menggunakan ensemble model MLP dengan memperhatikan 3 fitur bitcoin (harga, volume, biaya transaksi) dengan akurasi yang diperoleh diantara rentang 58% - 63% secara konsisten. Thearasak Phaladisailoed, dkk [2] membandingkan sejumlah algoritma pembelajaran mesin seperti LSTM, GRU, Theil-Sen dan Huber regresi dalam memprediksi harga bitcoin. Pada [3] dilakukan deteksi kecurangan pada transaksi bitcoin. Siddhi Velankar, dkk [4] melakukan prediksi terhadap harga bitcoin dengan menggunakan bayessian regression dan GLM atau random forest. Pada [5] dilakukan penelitian mengenai bitcoin untuk mengetahui jumlah alamat bitcoin dan alamat bitcoin berdasarkan karakteristik dari alamat deposit. Pada [8] telah dilakukan penelitian mengenai time series forecasting untuk bitcoin menggunakan data observasi short-term untuk 7 hari dan mid-term untuk 30-90 hari. Yang Li, dkk [16] bahkan melakukan penelitian yang menginvestigasi efek dari parameter harga bitcoin yang kedepannya dapat digunakan pada trading secara *real time*. Selanjutnya, R. Habibi [19] pada hasil penelitian menyatakan bahwa algoritma random forest memberikan akurasi terbaik pada saat memprediksi bitcoin dibandingkan dengan algoritma *decision tree* yang lain. Dan V. Derbentsev [20] pada bulan January 2021, melakukan riset untuk memprediksi harga *cryptocurrencies* menggunakan metode *ensemble learning*, dimana dinyatakan bahwa diperlukan *predictive power* dengan menggunakan penambahan *features*. Oleh karena itu, permasalahan prediksi bitcoin menjadi penelitian yang marak dilakukan sampai dengan saat ini untuk dipelajari lebih lanjut, khususnya merujuk kepada hasil penelitian yang telah dipublikasikan pada Saadah [12] tahun 2020 mengenai keterhubungan prediksi *cryptocurrency* terutama bitcoin terhadap kondisi stabilitas keuangan di Indonesia.

Mengacu kepada permasalahan di atas, maka pada penelitian ini dilakukan observasi mengenai prediksi harga bitcoin dengan memperhatikan fitur-fitur atau parameter apa yang paling mempengaruhi nilai harga bitcoin, terutama di masa awal pandemik yakni menggunakan dataset acak pada tahun 2019 dengan menggunakan metode Random Forest. Kemudian melihat nilai

kesalahan sistem prediksi dengan mengukur akurasi prediksi menggunakan *Mean Absolute Percentage Error* (MAPE).



Gambar 1. Fluktuasi Bitcoin Selama 30 Tahun [16]

2. Metode Penelitian

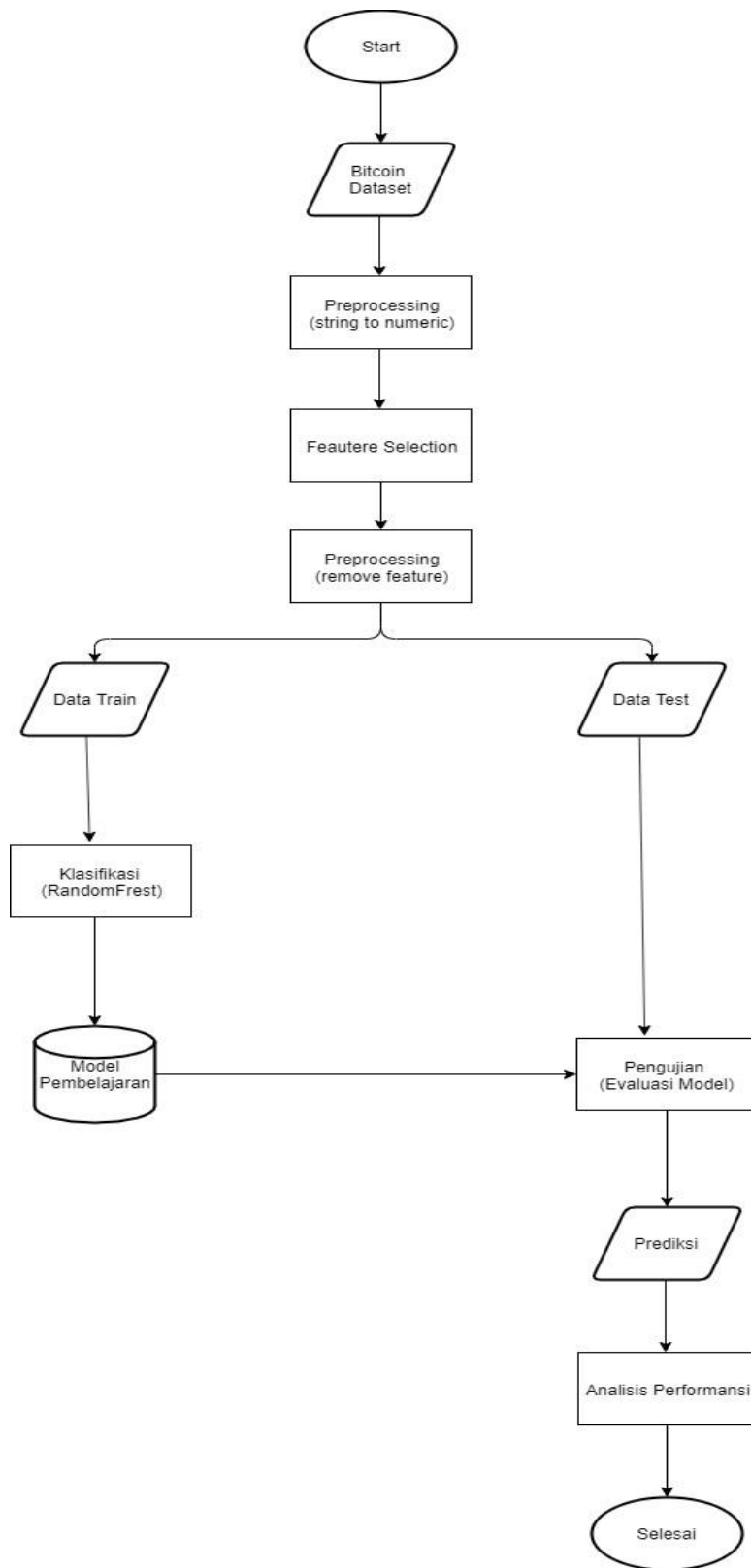
Pada penelitian ini terdapat 4 tahap yang dilakukan, yaitu pengumpulan dataset, preprocessing, feature selection, pembuatan model latih, pengujian dan pengukuran performansi sistem. Alur pembuatan sistem prediksi harga bitcoin diilustrasikan pada gambar 2.

2.1 Preprocessing Data

Dataset yang digunakan pada penelitian ini adalah data bitcoin yang didapatkan dari laman <https://investing.com/crypto/bitcoin/historical-data> dengan rentang waktu dari 1 Januari 2019 – 13 Oktober 2019 dengan fitur *Date*, *Price*, *Open*, *High*, *Low*, *Volume*, dan *Change %*. *Date* merupakan tanggal transaksi bitcoin. *Price* adalah harga bitcoin pada tanggal terkait. *Open* adalah harga awal bitcoin pada tanggal terkait. *High* adalah harga tertinggi bitcoin pada pada tanggal terkait. *Low* adalah harga terendah bitcoin pada tanggal terkait. *Volume* adalah total transaksi perdagangan bitcoin pada tanggal terkait. *Change %* adalah perubahan nilai harga bitcoin pada tanggal terkait terhadap harga bitcoin pada tanggal sebelumnya dalam bentuk persen. Tabel 1 di bawah ini menunjukkan sampling dari data set yang digunakan.

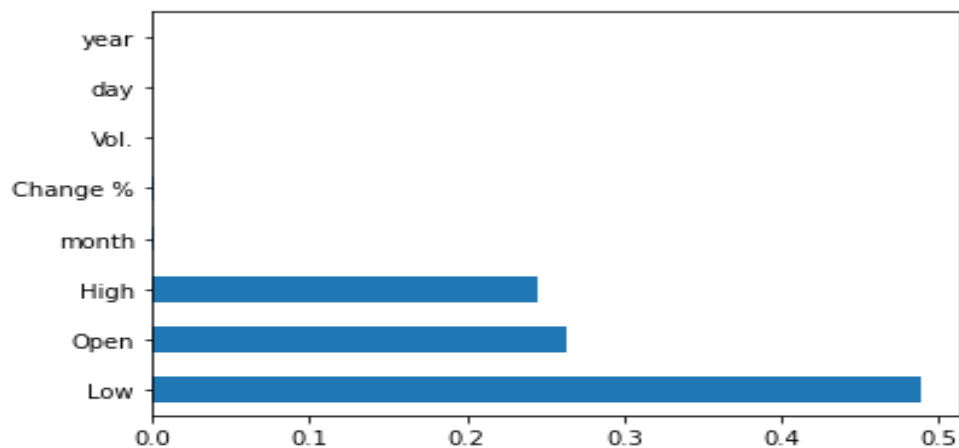
Tabel 1. Sampling Data Set

Date	Price	Open	High	Low	Vol.	Change %
13-Oct-19	8396.7	8304.3	8435.9	8289.1	374630	0.01
12-Oct-19	8304.4	8267.6	8394.8	8260.9	424020	0
11-Oct-19	8267.8	8562.2	8751.1	8249.7	667860	-0.03
10-Oct-19	8562.3	8566.8	8638.5	8438.7	511460	0
9-Oct-19	8566.7	8180.1	8655.9	8122.7	636430	0.05
8-Oct-19	8182.9	8198.5	8323.1	8118.1	469040	0
7-Oct-19	8198.6	7882.6	8293.5	7797	604740	0.04
6-Oct-19	7881.9	8127.3	8145.5	7834.9	448800	-0.03
5-Oct-19	8127.3	8148	8185.1	8030.6	360920	0



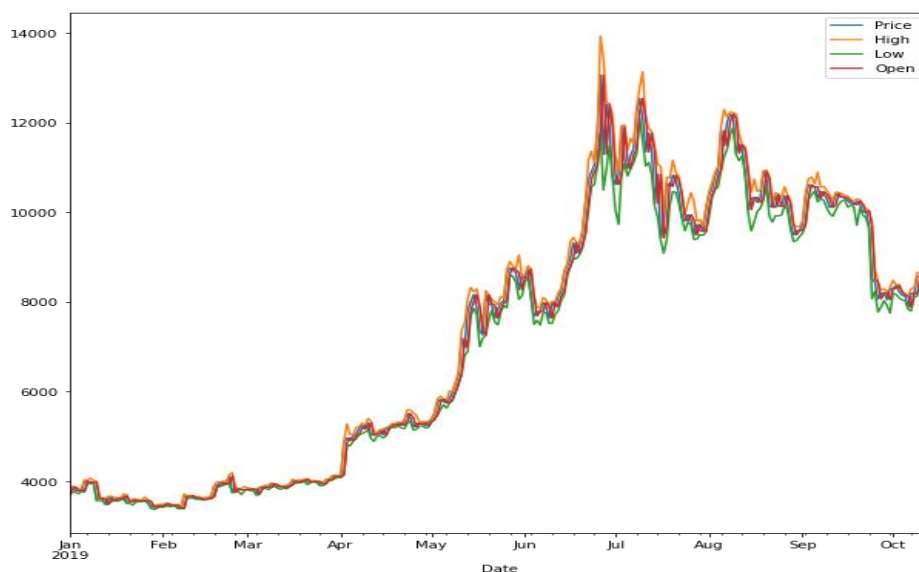
Gambar 2. Alur Sistem Prediksi Bitcoin menggunakan Random Forest

Pada sistem ini dilakukan dua kali tahap *preprocessing* pada waktu yang berbeda. Tahap *Preprocessing* pertama adalah mengubah *value Date* yang bertipe string menjadi *numeric*. *Value date* diubah menjadi *numeric* karena *random forest* hanya dapat menerima data input *numeric*. Kemudian tahap *preprocessing* kedua akan membentuk tiga kolom atribut baru, yaitu *year*, *month*, *day* dan membuang kolom atribut *Date*. Selanjutnya pada proses *feature selection* akan di observasi mengenai fitur atau atribut mana yang berpengaruh dan tidak berpengaruh terhadap harga bitcoin. Berdasarkan dataset yang dimiliki, diperoleh atribut *low*, *high* dan *open* yang merupakan feature tertinggi dari hasil *feature selection*. Adapun hasil dari proses ini ditampilkan pada gambar 3.



Gambar 3. Hasil *Feature Selection*

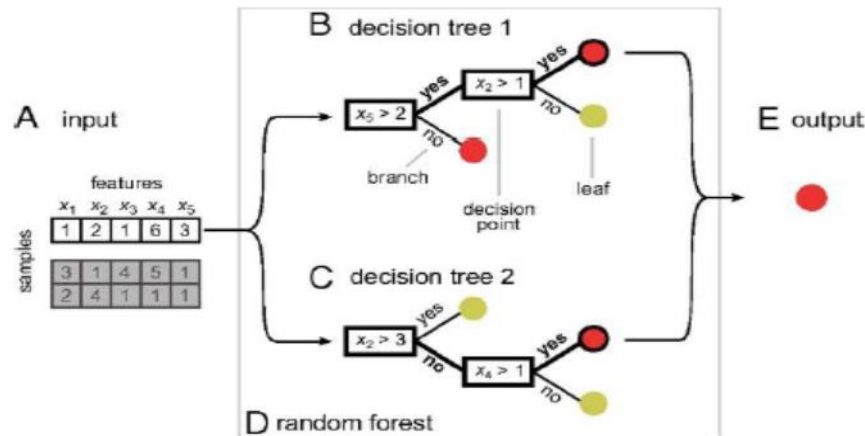
Gambar 4 adalah grafik pergerakan nilai harga bitcoin dari Januari 2019 hingga 13 Oktober 2019. Grafik tersebut menggambarkan bahwa nilai harga bitcoin bersifat fluktuatif, dengan tingkat naik atau turun yang *volatile*. Terlihat pada grafik bahwa beberapa bulan nilai harga bitcoin mengalami kenaikan yang sangat drastis, seperti pada bulan April, Mei, dan Juni. Kemudian mengalami penurunan yang signifikan pada bulan September.



Gambar 4. Grafik Pergerakan Nilai Harga Bitcoin

2.2 Random Forest

Random forest merupakan algoritma pembelajaran mesin supervised yang melakukan konsep *decision-tree* berulang kali sehingga solah membentuk suatu *forest* atau hutan. Di bawah ini berikut ilustrasi mengenai metode tersebut.



Gambar 5. Random Forest Method [18]

Random forest adalah metode machine learning yang menggunakan konsep supervised dalam membangun kelas classifier. Algoritma ini mengkombinasikan prediksi berdasarkan *multiple decision tree*. Menggunakan pendekatan ini diperoleh model pembelajaran dari pola atau trend dataset bitcoin. Untuk kemudian akan dijadikan masukan untuk fase pengujian untuk mengukur performansi sistem dalam memprediksi.

2.3 Pengujian dan Pengukuran Performansi Sistem

Pengujian sistem dilakukan untuk mengetahui pemodelan *training* yang telah diperoleh memberikan performansi yang bagus atau tidak. Pada penelitian ini dilakukan enam pengujian dengan perbandingan data latih dan data uji berbeda. Pengujian ini dilakukan untuk mengukur kemampuan random forest saat memprediksi harga bitcoin. Tiga pengujian pertama, cara pengambilan data dilakukan secara acak. Sedangkan tiga pengujian kedua, cara pengambilan data dilakukan secara tidak acak. Gambar 5 merupakan contoh data latih awal yang cara pengambilan data dilakukan secara acak.

Tabel 2. Sample Data Train dan Data Test Acak

Date	Price	Open	High	Low	Vol.	Change %
43751	8396.7	8304.3	8435.9	8289.1	374630	0.01
43750	8304.4	8267.6	8394.8	8260.9	424020	0
43738	8284.3	8048.9	8318.5	7749	600260	0.03
43737	8049.1	8208	8226.8	7922.1	418480	-0.02
43696	10910.7	10312.8	10923.9	10267.8	519010	0.06
43695	10312.7	10219.3	10494.9	10086.7	441110	0.01

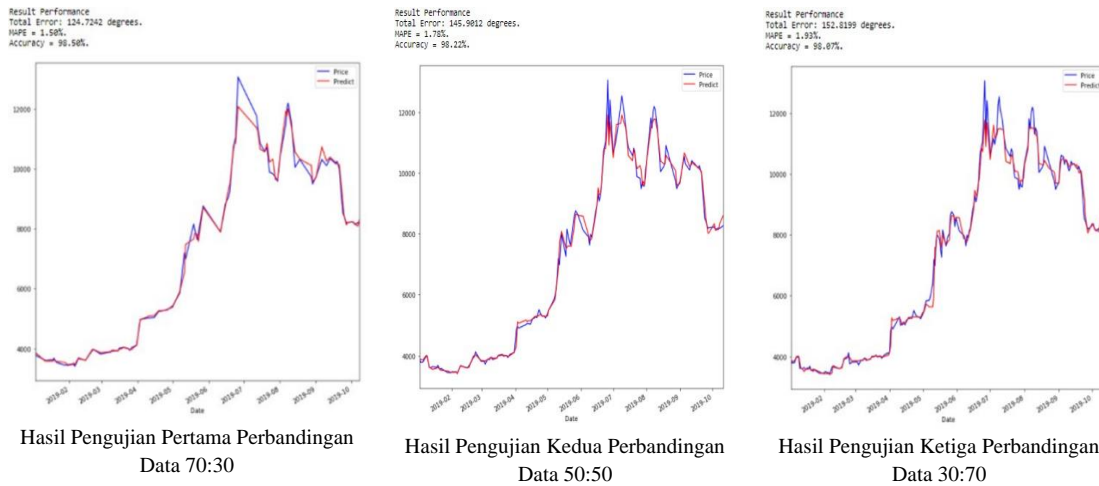
Untuk setiap pengujian dilakukan perhitungan performansi sistem berdasarkan nilai MAPE dan akurasi dengan persamaan yang digunakan seperti perumusan (1) dan (2).

$$MAPE = \frac{\sum total\ error}{Label\ Data\ Uji} \times 100\% \quad (1)$$

$$Akurasi = (100\% - MAPE) \quad (2)$$

3. Hasil Dan Pembahasan

Berdasarkan karakteristik data bitcoin bersifat *volatility*, maka pada penelitian ini dilakukan tiga pengujian dengan perbandingan jumlah data latih dan data uji yang berbeda.



Gambar 6. Hasil Pengujian Prediksi Harga Bitcoin Menggunakan Data Acak

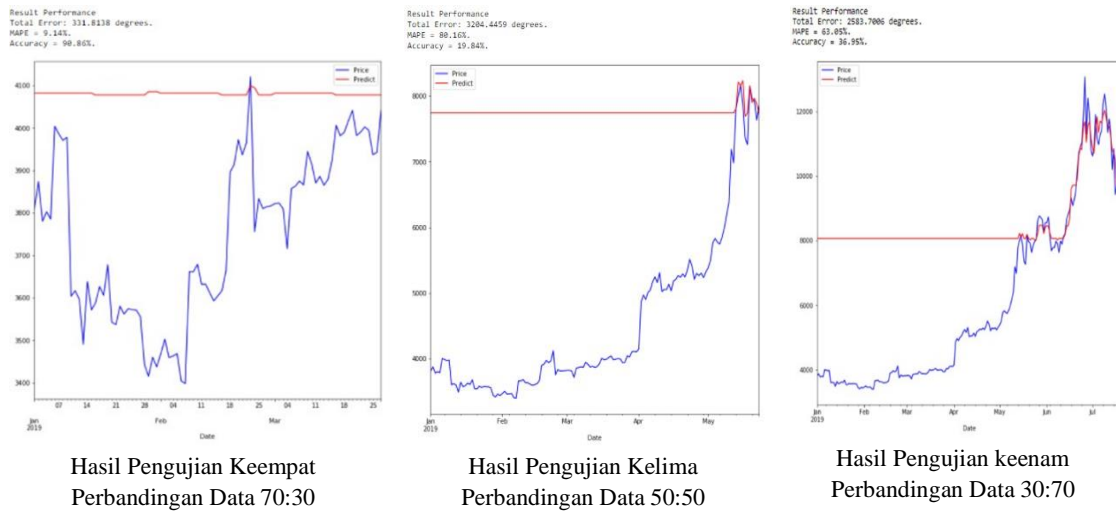
Berdasarkan pengujian yang telah dilakukan, hasil performansi dari prediksi harga bitcoin dengan algoritma random forest dipengaruhi oleh cara pengambilan data latih dan data uji. Pengujian dengan cara pengambilan data secara acak menghasilkan nilai MAPE yang lebih kecil dibandingkan dengan pengujian dengan cara pengambilan secara tidak acak. Atau dengan kata lain, prediksi bitcoin menggunakan data acak akan memberikan akurasi yang mendekati data actual yaitu dengan rentang akurasi berkisar 98%.

Berdasarkan hasil ini dapat dinyatakan bahwa karakteristik data bitcoin dengan tingkat fluktuatif yang tinggi lebih sesuai menggunakan pola prediksi dengan data acak, sehingga pembentukan multiple tree pada random forest mampu mengikuti trend atau pola data sesungguhnya.

Tabel 3. Pengukuran Performansi Dataset Acak

Skenario	Perbandingan Data	MAPE	Akurasi
Pengujian Pertama	70:30	1.50%	98.50%
Pengujian Kedua	50:50	1.76%	98.22%
Pengujian Ketiga	30:70	1.81%	98.19%

Untuk menghitung kompleksitas algoritma random forest dalam memprediksi harga bitcoin, maka akan dihitung berdasarkan awal desain umum sistem dengan cara menganalogikan variabel n adalah jumlah data latih yang digunakan pada pemodelan, m adalah jumlah fitur pada data, dan d adalah kedalaman *tree*, maka kompleksitas waktu adalah $O(nmd)$.



Gambar 7. Hasil Pengujian Prediksi Harga Bitcoin Menggunakan Data Tidak Acak

Sedangkan, pengujian menggunakan data tidak acak, memberikan hasil prediksi yang tidak dapat mengikuti pola data actual. Dapat dilihat pada gambar 7 bahwa meskipun menggunakan perbandingan data latih dan data uji yang berbeda-beda, sistem prediksi dengan random forest ini tetap tidak dapat fitting terhadap trend dataset, dimana rentang akurasi yang diperoleh sekitar 19% - 37%. Kegagalan random forest dalam memprediksi harga bitcoin pada skema ini menunjukkan bahwa pemilihan keputusan pada saat membangun tree secara berulang-ulang tidak dapat mewakili representasi dari pola data actual.

Tabel 4 Pengukuran Performansi Dataset Tidak Acak

Skenario	Perbandingan Data	MAPE	Akurasi
Pengujian Pertama	70:30	9.14%	90.86%
Pengujian Kedua	50:50	80.16%	19.84%
Pengujian Ketiga	30:70	63.05%	36.95%

4. Kesimpulan Dan Saran

Berdasarkan hasil penelitian, algoritma random forest dengan menggunakan atribut low, high dan price berhasil memprediksi harga bitcoin dengan performansi yang baik dengan nilai MAPE sebesar 1.50% atau dengan akurasi yang diperoleh sekitar 98% dengan menggunakan data acak. Perolehan akurasi untuk prediksi dengan data acak ini menunjukkan nilai yang setara yakni sekitar 95% - 98%. Sehingga mengacu kepada hasil prediksi ini dapat disimpulkan bahwa untuk memprediksi data yang memiliki karakteristik seperti bitcoin, yakni memiliki fluktuasi tinggi, random forest dikatakan berhasil memberikan fitting yang sesuai dengan data sesungguhnya. Akan tetapi, hasil ini berbanding terbalik pada saat memprediksi menggunakan data tidak acak.

Keberhasilan pada penelitian ini dapat menjadi *pilot project* untuk penelitian selanjutnya mengenai penelitian yang terkait dengan pergerakan nilai mata uang digital, seperti pada saat mengamati prediksi pergerakan global financial risk. Mengingat bahwa pembelajaran mengenai prediksi mata uang digital juga dapat memberikan efek untuk stabilitas nilai keuangan di sebuah negara secara tidak langsung, oleh sebab itu tidak menutup kemungkinan diperlukan penelitian yang menggabungkan algoritma ini dengan metode deep learning di masa yang akan datang.

Daftar Pustaka

- [1] E. Sin and L. Wang, 2017, Bitcoin Price Prediction Using Ensembles of Neural Network, 13th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov., pp. 666–671.
- [2] T. Phaladisailoed, July 2018, Machine Learning Models Comparison for Bitcoin Price Prediction, 10th Int. Conf. Inf. Technol. Electr. Eng., pp. 506– 511.
- [3] A. Viswam, 2017, An Efficient Bitcoin Fraud Detection in Social Media Networks, pp. 1–4.
- [4] S. Velankar, S. Valecha, S. Maji, and A. Bitcoin, 2018, Bitcoin Price Prediction using Machine Learning, pp. 144–147.
- [5] S. Hong and H. Kim, 2019, Analysis of Bitcoin Exchange Using Relationship of Transactions and Addresses, 21st Int. Conf. Adv. Commun. Technol., pp. 67–70.
- [6] M. S. Kumar, V. Soundarya, S. Kavitha, E. S. Keerthika, and E. Aswini, 2019, Credit Card Fraud Detection, 3rd Int. Conf. Comput. Commun. Technol., pp. 149–153.
- [7] B. J. Samajpati and S. D. Degadwala, 2016, “Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier,” no. 2013, pp. 1015–1019.
- [8] Mudassir M, et al. 4 July 2020. Time-series forecasting of Bitcoin prices using high dimensional features: a machine learning approach. Springer-Neural Computing and Applications <https://doi.org/10.1007/s00521-020-05129-6>
- [9] Fiadhi J, et al, 20 April 2020, Bitcoin Price Prediction using ARIMA Model, <https://www.researchgate.net/publication/328989226> DOI: 10.36227/techrxiv.12098067
- [10] Dharminder Singh Virk, 2016, Prediction of Bitcoin Price using Data Mining, National College of Ireland, School of Computing.
- [11] Ahmed I et al, January 2021, Predicting market movement direction for bitcoin: A comparison of time series modeling methods, Elsevier Computer & Electrical Engineering, <https://doi.org/10.1016/j.compeleceng.2020.106905>.
- [12] S Saadah, A Whafa, August 2020, Monitoring Financial Stability Based on Prediction of Cryptocurrencies Price Using Intelligent Algorithm. 2020 International Conference on Data Science and Its Applications (ICoDSA), <https://ieeexplore.ieee.org/document/9212968>.
- [13] Lekkala S.R and Siramya P, April 2020, A Research on Bitcoin Price Prediction Using Machine Learning Algorithm, International Journal of Scientific and Technology Resarch Volume 9.
- [14] Madan I et al, Automated Bitcoin Trading via Machine Learning Algorithm, Research from Department of Computer Science, Stanford University.
- [15] Sabah A and Ansari M, April 2020, Bitcoin Price Prediction Using ARIMA Model, [10.36227/techrxiv.12098067](https://doi.org/10.36227/techrxiv.12098067).
- [16] Yang Li, Zibin Zeng and Dai H.N, July 2020, Enhancing Bitcoin Price Fluctuation Prediction Using Attentive LSTM and Embedding Network.
- [17] Mudassir M et al, June 2020, Time-Series Forecasting of Bitcoin Prices Using High Dimensional Features: Machine Learning Approach, Springer Neural Computing and Applications.
- [18] Danielle Denisko and Michael M. Hoffman, 2018, Classification and interaction in random forests. PNAS February 20, 2018 115 (8) 1690-1692; first published February 12, 2018; <https://doi.org/10.1073/pnas.1800256115>.
- [19] Naghib A and Habibi R, December 2020, Crypto-Currency Price Prediction with Decision Tree Based Regressions Approach, Journal of Algorithms and Computation.
- [20] V. Derbentsev et al, January 2021, Comparative Performance of Machine Learning Ensemble Algorithms for Forecasting Cryptocurrency Prices, International Journal of Engineering.