

## IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK MEMREDIKSI HARGA SAHAM MENGGUNAKAN INDIKATOR EXPONENTIAL MOVING AVERAGE 10 PERIODE HARIAN

Welman Apridius<sup>1\*</sup>, Desiyanna Lasut<sup>2</sup>

<sup>1,2</sup> Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Buddhi Dharma

\*Corresponding Author, email: [desiyanna.lasut@ubd.ac.id](mailto:desiyanna.lasut@ubd.ac.id)

### ABSTRAK

Fluktuasi harga saham yang dinamis dan sulit diprediksi secara manual menjadi tantangan dalam pengambilan keputusan investasi yang tepat. Oleh karena itu, diperlukan pendekatan berbasis data yang mampu memberikan prediksi harga secara lebih akurat. Penelitian ini bertujuan untuk mengembangkan aplikasi prediksi harga saham menggunakan algoritma Random Forest dengan memanfaatkan data historis dari 10 perusahaan teratas dalam indeks LQ45. Data yang digunakan diperoleh dari API Yahoo Finance dan difokuskan pada indikator teknikal *Exponential Moving Average* (EMA) 10 periode harian. Proses praproses data meliputi normalisasi, visualisasi tren, dan pembersihan data. Selanjutnya, model Random Forest diterapkan untuk memprediksi harga saham satu hari ke depan. Algoritma ini dipilih karena kemampuannya dalam menangani hubungan non-linear dan memberikan hasil prediksi yang stabil. Evaluasi performa model dilakukan menggunakan metrik *Mean Absolute Percentage Error* (MAPE) untuk mengukur tingkat akurasi. Hasil penelitian menunjukkan bahwa model mampu menghasilkan prediksi harga dengan tingkat kesalahan rata-rata yang rendah. Aplikasi ini juga dilengkapi dengan fitur sinyal beli atau jual berdasarkan hasil prediksi, sehingga dapat membantu investor dalam menyusun strategi investasi yang lebih informatif. Dengan demikian, aplikasi ini diharapkan menjadi alat bantu yang efektif bagi investor pemula maupun profesional dalam pengambilan keputusan investasi berbasis data.

**Kata kunci:** *Exponential Moving Average* (EMA), Prediksi Harga Saham, *Random Forest*, *Website*, *Yahoo Finance*.

### I. PENDAHULUAN

Investasi di pasar saham menawarkan potensi keuntungan besar, namun diiringi volatilitas harga yang signifikan. Fluktuasi ini dipengaruhi oleh faktor makroekonomi, kebijakan pemerintah, dan sentimen pasar, menjadikan prediksi harga saham yang kompleks (Khandagale et al., 2023; Sudhakar Dharrao et al., 2023). Indeks Harga Saham Gabungan (IHSG) sebagai cerminan kondisi ekonomi nasional turut sensitif terhadap peristiwa global seperti pandemi COVID-19 (Eko Waluyo et al., 2024; Siregar et al., 2022). Untuk mengantisipasi ketidakpastian ini, analisis teknikal menjadi pilihan, dengan indikator seperti *Exponential Moving Average* (EMA-10) yang efektif menangkap tren jangka pendek (Bastian et al.,

2021), serta indikator lain seperti Bollinger Bands, On-Balance Volume, dan MACD yang umum digunakan (Valiant et al., 2019). Pemodelan data yang akurat sangat krusial mengingat pergerakan harga saham yang dinamis (Julian & Pribadi, 2021).

Meskipun analisis teknikal bermanfaat, keterbatasannya dalam menangani pola non-linear mendorong adopsi pendekatan *machine learning*. Random Forest, khususnya, banyak digunakan karena kemampuannya menangkap hubungan non-linear dan memberikan prediksi yang stabil dan akurat (Iqbal Baihaqi et al., 2023; Khandagale et al., 2023). Algoritma ini terbukti efektif dalam memprediksi arah pergerakan saham (Bastian et al., 2021; Valiant et al., 2019) dan digunakan dalam prediksi *return* saham untuk optimasi portofolio (Novi & Saepudin, 2023). Selain Random Forest, metode Long Short-Term Memory (LSTM) juga menunjukkan performa baik dalam peramalan harga saham, terutama untuk data historis jangka panjang (Julian & Pribadi, 2021). Pemilihan model prediksi yang tepat memerlukan evaluasi performa yang akurat. Dalam konteks regresi, Mean Squared Error (MSE) adalah metrik kunci, di mana Random Forest secara otomatis menyertakan Weighted MSE yang mempertimbangkan kontribusi setiap pohon. Selain itu, Mean Absolute Percentage Error (MAPE) sangat relevan untuk evaluasi akurasi data riil karena menyajikan kesalahan dalam persentase, memudahkan interpretasi bagi investor (Chicco et al., 2021). Oleh karena itu, kombinasi metrik seperti Weighted MSE dan MAPE lazim digunakan untuk memastikan objektivitas dan relevansi hasil prediksi harga saham dengan praktik di lapangan (Fitri & Riana, 2022).

Dalam implementasi model, pengumpulan data historis adalah esensial. Pemanfaatan Application Programming Interface (API) seperti Yahoo Finance API mempermudah akuisisi data saham secara *real-time* dan otomatis, meningkatkan efisiensi pengembangan sistem prediksi (Karnyoto, 2018). Bahasa pemrograman Python menjadi pilihan utama dalam pengembangan model *machine learning* berkat kelengkapan pustaka dan komunitasnya yang luas (Raschka et al., 2020). Metodologi pengembangan sistem prediksi saham umumnya mengadopsi CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Kerangka kerja ini, yang mencakup tahapan pemahaman bisnis hingga *deployment*, merupakan standar *de-facto* dalam proyek *data mining* dan terbukti menghasilkan akurasi tinggi dalam

studi kasus seperti saham TLKM (Pambudi, 2023; Schröer et al., 2021). Penelitian ini mengintegrasikan analisis teknikal (EMA-10), kekuatan Random Forest, pengambilan data via API, dan pengembangan sistem berbasis Python, dengan tujuan utama menyediakan sistem prediksi saham berbasis *web* yang mudah diakses guna mendukung keputusan investasi berbasis data.

## II. METODOLOGI

Penelitian ini mengikuti pendekatan kuantitatif dengan metode data mining menggunakan kerangka kerja CRISP-DM (Cross-Industry Standard Process for Data Mining). Kerangka kerja ini dipilih karena merupakan standar industri yang sistematis dan fleksibel dalam membimbing seluruh tahapan data mining, mulai dari pemahaman bisnis hingga evaluasi dan deployment. CRISP-DM telah terbukti menjadi metode yang paling banyak digunakan dalam proyek data mining karena menyediakan kerangka kerja yang dapat disesuaikan dengan berbagai domain dan kebutuhan bisnis (Schröer et al., 2021).



Gambar 1. Siklus CRISP-DM

## III. HASIL DAN PEMBAHASAN

### 3.1 Business Understanding

Tahap awal ini difokuskan pada pemahaman kebutuhan bisnis, yaitu pengembangan model prediksi harga saham berbasis web menggunakan algoritma Random Forest. Target utama adalah saham-saham indeks LQ45 di Bursa Efek Indonesia, dengan fokus pada prediksi jangka pendek menggunakan indikator Exponential Moving Average (EMA) 10 periode harian.

### 3.2 Data Understanding

Data yang digunakan berupa harga penutupan saham harian yang diambil secara otomatis melalui Yahoo Finance API. Data mencakup 10 perusahaan dari indeks LQ45 selama dua tahun, mulai 31 Oktober 2022 hingga 31 Oktober 2024.

**Tabel 1. 10 Data Saham ADRO**

Date	Close	High	Low	Open	Volume	EMA_10	Return
2022-10-31 00:00:00	1732,2	1732,2	1706,09	1714,79	76758600	1707,27	0,018
2022-11-01 00:00:00	1632,1	1714,79	1619,04	1714,79	175308100	1693,61	-0,058
2022-11-02 00:00:00	1627,75	1662,56	1605,98	1632,1	101047500	1681,63	-0,003
2022-11-03 00:00:00	1632,1	1658,21	1623,39	1627,75	57965500	1672,62	0,003
2022-11-04 00:00:00	1632,1	1649,51	1614,69	1623,39	84299700	1665,26	0
2022-11-07 00:00:00	1666,92	1666,92	1623,39	1632,1	74309100	1665,56	0,021
2022-11-08 00:00:00	1636,45	1653,86	1623,39	1653,86	71681200	1660,27	-0,018
2022-11-09 00:00:00	1584,22	1619,04	1579,87	1605,98	117969800	1646,44	-0,032
2022-11-10 00:00:00	1514,59	1549,41	1501,53	1523,29	153812400	1622,47	-0,044
2022-11-11 00:00:00	1518,94	1540,7	1497,18	1523,29	96575300	1603,64	0,003

### 3.3 Data Preparation

Tahapan ini melibatkan proses pembersihan data yang berfungsi untuk menghapus data NaN (*Not a Number*) dan menghasilkan keseimbangan data.

**Tabel 2. Kode Pembersihan Data**

Baris Kode	Penjelasan
<code>data = pd.read_excel(file_path)</code>	Membaca file Excel yang berisi data saham berdasarkan path file yang diberikan.
<code>data = data[['Close', 'EMA_10', 'Return']].dropna()</code>	Memilih hanya kolom Close, EMA_10, dan Return, lalu menghapus baris yang memiliki nilai kosong.

### 3.4 Modelling Phase

Ini adalah tahapan inti dari proses penelitian, tahapan ini melibatkan pemilihan fitur yang telah diproses pada tahap sebelumnya, split data untuk

pelatihan model menggunakan fitur yang sudah dipilih serta pembangunan model menggunakan algoritma Random Forest Regressor.

**Tabel 3. Kode Pemilihan Fitur**

Baris Kode	Penjelasan
<code>X = data[['Close', 'EMA_10', 'Return']]</code>	Menentukan fitur yang digunakan untuk pelatihan model, yaitu Close, EMA_10, dan Return.
<code>y = data['Close'].shift(-1).dropna()</code>	Menentukan target prediksi yaitu harga Close pada hari berikutnya dengan cara menggeser 1 baris ke atas.
<code>X = X.iloc[:-1]</code>	Menyesuaikan jumlah baris X agar selaras dengan y, karena y telah digeser ke depan satu periode.

**Tabel 4. Kode Split Data**

Baris Kode	Penjelasan
<code>X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)</code>	Membagi data menjadi data latih (80%) dan data uji (20%) secara acak namun konsisten menggunakan <code>random_state=42</code> .

**Tabel 5. Pembuatan Model dan Latih**

Baris Kode	Penjelasan
<code>model = RandomForestRegressor(n_estimators=100, random_state=42)</code>	Membuat model regresi menggunakan algoritma Random Forest dengan 100 pohon keputusan dan random state untuk konsistensi.
<code>model.fit(X_train, y_train)</code>	Melatih model menggunakan data latih (X_train dan y_train) agar dapat mempelajari pola dari data historis.

### 3.5 Evaluation

Evaluasi yang digunakan untuk mengukur keakuratan prediksi dalam proyek ini adalah Mean Absolute Percentage Error (MAPE).

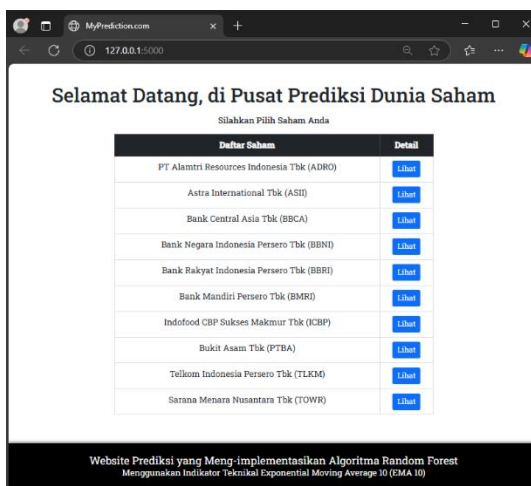
**Tabel 6. Menghitung MAPE**

Baris Kode	Penjelasan
<code>actual_values = data_clean['Close'][-10:].values</code>	Mengambil 10 nilai aktual terakhir dari kolom Close dan mengubahnya menjadi array untuk dibandingkan dengan hasil prediksi.
<code>predicted_values = model.predict(data_clean[['Close', 'EMA_10', 'Return']][-10:])</code>	Menghasilkan 10 prediksi nilai Close berdasarkan fitur Close, EMA_10, dan Return dari data terbaru.
<code>accuracy = calculate_accuracy(actual_values, predicted_values)</code>	Menghitung tingkat akurasi prediksi menggunakan fungsi <code>calculate_accuracy</code> ,

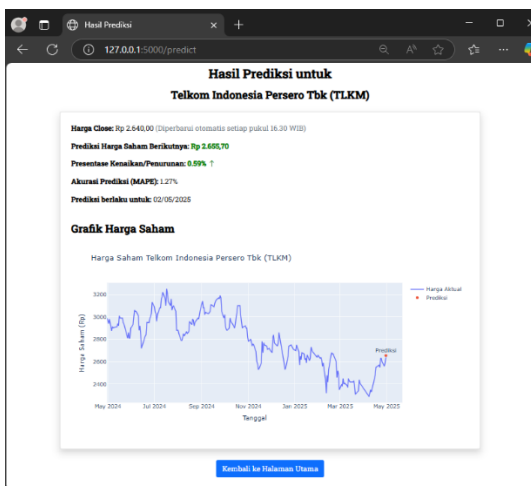
	biasanya berupa MAPE (Mean Absolute Percentage Error).
--	--

### 3.6 Deployment

Setelah melalui seluruh tahapan dalam metode CRISP-DM, model prediksi harga saham berhasil dikembangkan dan diimplementasikan dalam bentuk aplikasi berbasis web. Berikut hasil implementasinya.



Gambar 2. Halaman Dashboard



Gambar 3. Halaman Hasil Prediksi

## IV. SIMPULAN

Penelitian ini berhasil mengembangkan sebuah aplikasi prediksi harga saham berbasis web dengan memanfaatkan algoritma *Random Forest* dan indikator teknikal *Exponential Moving Average (EMA) 10*. Implementasi model menunjukkan bahwa pendekatan *Random Forest* dapat digunakan secara efektif

untuk memprediksi harga saham jangka pendek. Aplikasi ini juga menyediakan fitur visualisasi tren harga serta sinyal beli/jual, yang dapat membantu investor dalam menganalisis kondisi pasar secara lebih objektif. Dengan adanya sistem ini, pengguna, baik pemula maupun profesional, dapat memperoleh informasi prediksi saham secara cepat, akurat, dan praktis melalui platform berbasis web.

## DAFTAR PUSTAKA

- Bastian, M. E., Rahayudi, B., & Ratnawati, D. E. (2021). Prediksi Trend Harga Saham Jangka Pendek berdasarkan Fitur Technical Analysis dengan menggunakan Algoritma Random Forest. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(10), 4536–4542. <http://j-ptiik.ub.ac.id>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- Eko Waluyo, D., Wikan Kinasih, H., Paramita, C., Pergiwati, D., Adi Rafrastara, F., Nohan, R., Studi Akuntansi, P., Ekonomi dan Bisnis, F., & Dian Nuswantoro, U. (2024). Implementasi Algoritma Regresi pada Machine Learning untuk Prediksi Indeks Harga Saham Gabungan. *Universitas Dian Nuswantoro, Semarang Jln. Imam Bonjol*, 9(1), 12–17.
- Fitri, E., & Riana, D. (2022). Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer Perceptron. *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi*, 6(1), 69–78. <https://doi.org/10.46880/jmika.vol6no1.pp69-78>
- Iqbal Baihaqi, M., Syaripudin, A., & Agung Nugroho, F. (2023). Implementasi Algoritma Random Forest Pada Prediksi Harga Saham Berdasarkan Data Historis. *Jubitek: JURNAL BIG DATA DAN TEKNOLOGI INFORMASI*, 1, 42–51.
- Julian, R., & Pribadi, M. R. (2021). Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM). *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(3),

- 1570–1580. <https://doi.org/10.35957/jatisi.v8i3.1159>
- Karnyoto, A. S. (2018). Mengunduh Data Saham Yahoo Finance Menggunakan YahooFinanceAPI. *Journal Dynamic Saint*, 3(2). <https://doi.org/10.47178/dynamicsaint.v3i2.312>
- Khandagale, H. P., Patil, R., Patil, S., & Bhosale, D. (2023). Predicting Stock Prices With Machine Learning Using Comparative Analysis of Random Forest Algorithm. *International Journal of Engineering Applied Sciences and Technology*, 8(6), 60–68. <https://doi.org/10.33564/ijeast.2023.v08i06.008>
- Novi, M. S., & Saepudin, D. (2023). Prediksi Return Saham dengan Metode Random Forest dan Penerapannya untuk Seleksi Portofolio. *E-Proceeding of Engineering*, 10(5), 4996–5001.
- Pambudi, A. (2023). Penerapan Crisp-Dm Menggunakan Mlr K-Fold Pada Data Saham Pt. Telkom Indonesia (Persero) Tbk (Tlkm) (Studi Kasus: Bursa Efek Indonesia Tahun 2015-2022). *Jurnal Data Mining Dan Sistem Informasi*, 4(1), 1. <https://doi.org/10.33365/jdmsi.v4i1.2462>
- Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information (Switzerland)*, 11(4). <https://doi.org/10.3390/info11040193>
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181(2019), 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Siregar, B., Anthon Pangruruk, F., & Widjaja, P. A. (2022). Comparison of Various Forecasting Models for the Jakarta Stock Exchange Composite (JKSE) During the Covid-19 Pandemic. *Jurnal Multidisiplin Madani (MUDIMA)*, 2(2), 1035–1046. <https://journal.yp3a.org/index.php/mudima/index>
- Sudhakar Dharrao, D., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Krishna Bhat, S. (2023). Citation: Sonkavde, Gaurang. <https://doi.org/10.3390/ijfs11030094>
- Valiant, K., Lukito, Y., & Santosa, R. G. (2019). Sistem Prediksi Harga Saham LQ45 Dengan Random Forest Classifier. *Jurnal Terapan Teknologi Informasi*, 3(2), 127–136. <https://doi.org/10.21460/jutei.2019.32.187>