

Implementasi Sistem Prediksi Saham Real-Time dengan *Integrasi Yahoo Finance API* dan *Machine Learning* di *Google Colab*

Muhammmad faiq abdi¹, Yonhendri²

^{1,2}Stikom Muhammadiyah batam

abdifaiq9@gmail.com¹, yonhendri@gmail.com²

ABSTRACT

The issue of stock price prediction is a critical topic in the financial world, where accurate predictions are essential for making better investment decisions. This research develops a real-time stock price prediction system by integrating the Yahoo Finance API and machine learning algorithms, executed on the Google Colab platform. This system allows for direct retrieval of stock market data from Yahoo Finance to analyze data patterns and generate more accurate stock price predictions. The methods used include collecting historical and real-time data from the Yahoo Finance API, data preprocessing, training models using Long Short Term Memory (LSTM), validating the model with K-Fold Cross Validation, and evaluating performance using various standard metrics. The development and implementation of the model are carried out on Google Colab. The results show that the LSTM model can provide high-accuracy stock price predictions. This system significantly contributes to improving stock price prediction accuracy and helps investors make better decisions based on real-time data

Keywords : *Stock prediction, real-time system, Yahoo Finance API, machine learning, Google Colab.*

ABSTRAK

Masalah prediksi harga saham merupakan topik penting dalam dunia keuangan, di mana akurasi prediksi sangat diperlukan untuk pengambilan keputusan investasi yang lebih baik. Penelitian ini mengembangkan sistem prediksi harga saham real-time dengan integrasi *Yahoo Finance API* dan algoritma *Machine Learning* yang dijalankan di *Google Colab*. Sistem ini memungkinkan pengambilan data pasar saham secara langsung dari *Yahoo Finance* untuk menganalisis pola data dan menghasilkan prediksi harga saham yang lebih akurat. Metode yang digunakan meliputi pengumpulan data historis dan real-time dari *Yahoo Finance API*, *preprocessing data*, pelatihan model menggunakan *Long Short Term Memory (LSTM)*, validasi model dengan *K-Fold Cross Validation*, dan evaluasi kinerja menggunakan berbagai metrik standar. Pengembangan dan implementasi model dilakukan di *Google Colab*. Hasil penelitian menunjukkan bahwa model LSTM mampu memberikan prediksi harga saham dengan akurasi tinggi. Sistem ini memberikan kontribusi signifikan dalam meningkatkan akurasi prediksi harga saham dan membantu investor membuat keputusan yang lebih baik berdasarkan data real-time.

Kata kunci : *Prediksi saham, sistem real-time, Yahoo Finance API, machine learning, Google Colab*

PENDAHULUAN

Dalam era digital yang semakin berkembang pesat, prediksi harga saham menjadi topik yang sangat penting bagi investor dalam pengambilan keputusan investasi di pasar modal. Peran teknologi memiliki peran penting dalam mengembangkan sistem prediksi yang lebih akurat dan efisien.

Salah satu inovasi terbaru adalah implementasi sistem prediksi saham real-time yang mengintegrasikan *Yahoo Finance API* dengan teknologi *Machine Learning* melalui

platform Google Colab. Sistem ini bertujuan untuk memberikan prediksi harga saham yang lebih akurat dan tepat waktu, sehingga investor dapat membuat keputusan yang lebih baik berdasarkan data yang terkini dan relevan.

Penelitian terkait telah menunjukkan bahwa penggunaan algoritma Deep Learning, khususnya Long Short Term Memory (LSTM), dapat membantu dalam memprediksi harga saham dengan lebih baik dibandingkan metode tradisional. LSTM adalah salah satu jenis Recurrent Neural Network (RNN) yang mampu menangani data berurutan dan mengatasi masalah fluktuasi harga saham [1]. Selain itu, integrasi dengan Yahoo Finance API memungkinkan akses ke data saham secara real-time, yang sangat penting untuk analisis dan prediksi yang akurat.

Pengembangan model prediksi saham menggunakan Google Colab memiliki beberapa keuntungan, pengembangan dapat dilakukan secara efisien dan cepat tanpa memerlukan perangkat keras yang canggih. Google Colab juga menyediakan akses mudah ke sumber daya komputasi berbasis cloud, yang memungkinkan pengembang untuk melakukan eksperimen dan pengujian model dengan cepat [1].

Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan sistem prediksi saham secara real-time dengan mengintegrasikan Yahoo Finance API dengan teknologi Machine Learning pada Google Colab. Sistem ini diharapkan dapat memberikan kontribusi yang signifikan dalam bidang keuangan dan investasi, serta membantu investor dalam membuat keputusan yang lebih baik berdasarkan prediksi harga saham yang akurat dan real-time.

METODE PENELITIAN

Penelitian ini menggunakan metode yang meliputi pengumpulan data historis dan real-time melalui *Yahoo Finance API*, dilanjutkan dengan tahap *preprocessing data*, pembangunan model menggunakan *Long Short Term Memory (LSTM)*, validasi model dengan *K-Fold Cross Validation*, dan evaluasi kinerja menggunakan berbagai metrik standar.

Pengumpulan Data

Pengumpulan data harga saham dilakukan melalui *Yahoo Finance API*, yang menyediakan akses ke data historis serta data *real-time* dari berbagai pasar saham di seluruh dunia. Penggunaan *Yahoo Finance API* memungkinkan peneliti untuk memperoleh data yang komprehensif dan relevan untuk analisis prediksi harga saham. API ini telah banyak digunakan dalam penelitian sebelumnya untuk mendapatkan data pasar yang akurat dan terkini [1][2].

Data yang akan diambil mencakup harga penutupan harian saham, volume perdagangan, dan informasi terkait lainnya. Proses pengambilan data dilakukan menggunakan pustaka *yfinance* dalam *Python*, yang memudahkan akses dan pengunduhan data dari *Yahoo Finance*. Dengan menggunakan *yfinance*, peneliti dapat mengunduh data dalam format *Data Frame* yang mudah untuk dianalisis dan diproses lebih lanjut [3].

Periode pengumpulan data mencakup rentang waktu 1 Januari 2018 hingga 1 Januari 2023. Rentang waktu ini dipilih untuk menganalisis tren harga saham dalam jangka

waktu yang cukup panjang, sehingga dapat memberikan wawasan yang lebih baik mengenai pola pergerakan harga dan faktor-faktor yang mempengaruhi harga saham [4].

Pemrosesan Data

Setelah proses pengunduhan data, langkah selanjutnya adalah melakukan pembersihan dan pra-pemrosesan data. Tahap ini mencakup penghapusan nilai yang hilang, normalisasi data menggunakan metode *Min-Max Scaling*, dan pembentukan *dataset* untuk pelatihan model. Proses ini memiliki peran penting dalam memastikan model yang dibangun dapat berfungsi dengan baik dan menghasilkan prediksi yang akurat [5].

Data yang telah dibersihkan akan dinormalisasi menggunakan metode *Min-Max Scaler*. Tahap normalisasi bertujuan untuk memastikan semua fitur berada dalam rentang skala yang sama, yang sangat penting untuk meningkatkan kinerja model *machine learning*. Melalui proses normalisasi, model akan lebih mudah untuk memahami pola dalam data tanpa terpengaruh oleh skala yang berbeda-beda [5].

Setelah proses normalisasi data selesai, *dataset* akan dibagi menjadi dua bagian: data latih (70%) dan data uji (30%). Pembagian ini dilakukan untuk mengevaluasi kinerja model secara akurat, dimana data latih dimanfaatkan untuk melatih model dan data uji digunakan untuk menguji kemampuan model setelah pelatihan selesai. Pembagian data membantu dalam memastikan model memiliki kemampuan generalisasi yang baik dan melakukan prediksi akurat pada data baru yang belum pernah diproses sebelumnya [6].

Pembangunan Model

Pembangunan model dimulai dengan pemilihan algoritma. *Long Short-Term Memory* (LSTM) dipilih untuk membangun model prediksi harga saham, karena kemampuannya dalam menangani data *time-series* dan kemampuannya dalam menangkap pola jangka panjang dalam data. LSTM adalah jenis jaringan saraf tiruan yang dirancang untuk memproses urutan data dengan cara yang mempertimbangkan konteks temporal [7].

Model LSTM akan dirancang dengan beberapa lapisan tersembunyi dan jumlah neuron yang dioptimalkan untuk mencapai keseimbangan antara *overfitting* dan *underfitting*. *Overfitting* terjadi ketika model terlalu sesuai dengan data latih sehingga tidak mampu menggeneralisasi pada data baru, sedangkan *underfitting* terjadi ketika model terlalu sederhana dan tidak mampu menangkap pola dalam data latih. Dengan pemilihan arsitektur yang tepat, model diharapkan memiliki performa yang baik dalam memprediksi harga saham [7].

Pelatihan Model

Model akan dilatih menggunakan *dataset* latih dengan parameter yang disesuaikan seperti *batch size*, jumlah *epoch*, dan fungsi aktivasi yang sesuai. *Batch size* menentukan jumlah sampel yang diproses sebelum model diperbarui, sedangkan *epoch* adalah jumlah iterasi pelatihan melalui seluruh *dataset*. Fungsi aktivasi membantu model untuk belajar pola non-linear dalam data [8].

Untuk mencegah *overfitting*, berbagai teknik optimasi akan diterapkan, termasuk *Early Stopping* dan *Regularization*. *Early Stopping* berfungsi untuk menghentikan pelatihan

model ketika performa pada data uji mulai menurun, sedangkan *Regularization* dapat menambahkan penalti pada kompleksitas model untuk mencegah model menjadi terlalu rumit [8].

Evaluasi Model

Kinerja model akan dievaluasi menggunakan metrik *Mean Squared Error* (MSE) dan *Mean Absolute Error* (MAE). MSE bertujuan mengukur rata-rata kuadrat kesalahan prediksi, sedangkan MAE untuk mengukur rata-rata kesalahan absolut. Kedua metrik ini memberikan pandangan yang komprehensif tentang seberapa baik model dapat memprediksi harga saham .

Validasi silang (*cross-validation*) juga akan dilakukan untuk memastikan bahwa model memiliki generalisasi yang baik pada data yang belum pernah dilihat sebelumnya. Dalam validasi silang, *dataset* dibagi menjadi beberapa *subset*, dan model dilatih dan diuji pada *subset* yang berbeda-beda untuk memastikan bahwa model dapat menangani variasi dalam data [7].

Implementasi

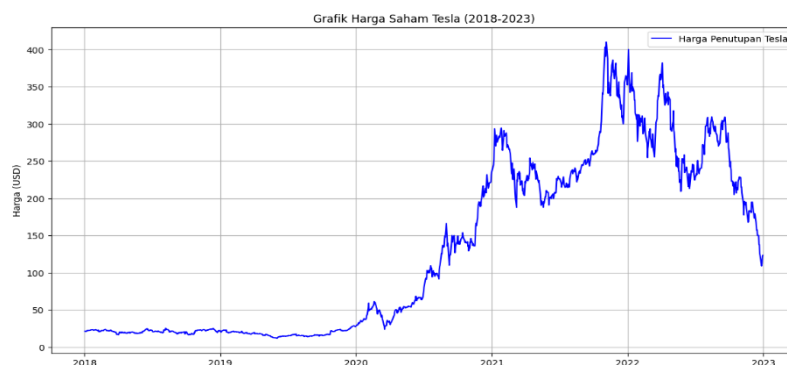
Model yang telah dilatih dan divalidasi akan diimplementasikan dalam sistem lokal yang dapat melakukan prediksi harga saham secara *real-time*. Prediksi ini akan dilakukan dengan menjalankan *script Python* di *Google Colab* untuk memanfaatkan kekuatan komputasi dan kemudahan akses [9].

Pengujian model dilakukan untuk memastikan bahwa semua fungsi bekerja dengan baik dan prediksi yang dihasilkan adalah akurat. Pengujian juga akan mencakup analisis hasil prediksi untuk memastikan kesesuaian dengan kondisi pasar saat ini [10].

HASIL DAN PEMBAHASAN

Hasil dan pelatihan model

Setelah melakukan pelatihan model menggunakan data historis harga saham Tesla dari tahun 2018 hingga 2023, model *Long Short Term Memory* (LSTM) dan *RandomForest* menunjukkan performa yang baik dalam memprediksi harga saham. Hasil pelatihan menunjukkan bahwa kedua model berhasil belajar pola harga saham dengan cukup akurat. Gambar 1 menunjukkan tren harga saham Tesla selama periode pelatihan dan prediksi.



Gambar 1. Tren harga saham Tesla selama periode pelatihan dan prediksi

Metrik Evaluasi

Tabel 1. Evaluasi MSE dan MAE

Model	MSE	MAE
LSTM	73325,258586	266,184550
<i>Random Forest</i>	76498,734178	270,855634

Hasil evaluasi menunjukkan bahwa model LSTM dan *Random Forest* berhasil memprediksi harga saham dengan akurasi yang baik. Tabel 1 menunjukkan bahwa model LSTM memiliki nilai MSE 73325,26 yang sedikit lebih rendah dibandingkan *Random Forest* yang mencapai 76498,73. Nilai MSE yang lebih rendah menunjukkan bahwa model memiliki kesalahan prediksi yang lebih kecil, sehingga LSTM lebih unggul dalam hal ini. Sementara itu, LSTM juga memiliki nilai MAE lebih kecil daripada *Random Forest* yaitu sebesar 266,18 (Tabel 1). Ini menandakan bahwa rata-rata kesalahan absolut pada prediksi yang dihasilkan oleh LSTM lebih rendah, yang berarti konsistensi hasil prediksi dari model lebih baik.

Tabel 2. Persentase Keberhasilan

Nama Model	Persentase Keberhasilan (%)
LSTM	93,292301%
<i>Random Forest</i>	94,771736%

Tabel 2 menunjukkan persentase keberhasilan model. *Random Forest* memiliki tingkat keberhasilan sedikit lebih besar yaitu 93,29%, dibandingkan LSTM yang memiliki tingkat keberhasilan 94,77%. Meskipun LSTM memiliki tingkat keberhasilan yang lebih rendah, namun perbedaannya relatif kecil. Sehingga dapat disimpulkan bahwa kedua model memiliki tingkat performa yang sebanding.

Diskusi

Berdasarkan hasil evaluasi, LSTM maupun *Random Forest* dapat bekerja dengan cukup akurat dalam memprediksi harga saham. LSTM memiliki nilai MSE dan MAE lebih tinggi dan lebih unggul dalam menangani kesalahan prediksi dibandingkan *Random Forest*. Sedangkan *Random Forest* memiliki persentase keberhasilan lebih tinggi dibandingkan LSTM, ini menunjukkan *Random Forest* dapat memberikan prediksi yang lebih akurat pada beberapa kondisi tertentu.

Namun, prediksi harga saham selalu memiliki tingkat ketidakpastian karena dipengaruhi oleh berbagai faktor eksternal. Selain itu, model ini dapat ditingkatkan dengan menambahkan fitur. Oleh karena itu, untuk dapat meningkatkan performa model, fitur tambahan seperti volume perdagangan, indeks pasar, atau data fundamental lainnya dapat ditambahkan ke dalam model.

KESIMPULAN DAN SARAN

Terdapat beberapa keterbatasan model dalam memprediksi harga saham. Pertama, Model LSTM dan *RandomForest* sangat bergantung pada data historis yang tersedia. Jika data historis tidak mencakup semua kondisi pasar, model mungkin tidak dapat memprediksi dengan akurat pada kondisi pasar baru.

Kedua, model memiliki kemungkinan mengalami *overfitting* pada data latih, sehingga model tidak dapat melakukan generalisasi dengan baik pada data uji. Teknik seperti *Early Stopping* dan *Regularization* dapat digunakan untuk mencegah *overfitting*.

Ketiga, model LSTM yang kompleks dapat memerlukan waktu komputasi yang lebih lama dan memerlukan sumber daya yang lebih besar. Hal ini dapat menjadi tantangan dalam mengolah data dalam jumlah besar atau melakukan prediksi secara *real-time*.

Berdasarkan hasil evaluasi dan keterbatasan yang dimiliki, model dapat ditingkatkan untuk meningkatkan akurasi dan kinerjanya.

DAFTAR PUSTAKA

- A. Agusta, I. Ernawati, dan A. Muliawati, "Prediksi Pergerakan Harga Saham Pada Sektor Farmasi Menggunakan Algoritma Long Short-Term Memory," *Jurnal Ilmu Komputer dan Informasi*, vol. 17, no. 2, pp. 3651, 2024. [Online]. Tersedia: <https://dx.doi.org/10.52958/iftk.v17i2.3651>
- A. G. Aleyusta, "LONG SHORT TERM MEMORY (VAR-LSTM) PADA PERAMALAN," *Digilib Unila*, 2023. [Online]. Tersedia: <http://digilib.unila.ac.id/82155/3/3.%20SKRIPSI%20TANPA%20BAB%20PEMBAHASAN%20-%20A.%20GILANG%20ALEYUSTA%20SAVADA.pdf>
- A. Rosyd et al., "Penerapan Metode Long Short Term Memory (LSTM) dalam Memprediksi Harga Saham PT Bank Central Asia," *STMIK IKMI Cirebon*, 2024. [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/article/download/8440/5012/>.
- Bhowmick, A., Rahman, A., & Rahman, R. M. (2019). Performance Analysis of Different Recurrent Neural Network Architectures and Classical Statistical Model for Financial Forecasting: A Case Study on Dhaka Stock Exchange (pp. 277–286). https://doi.org/10.1007/978-3-030-19810-7_27
- C. A. F. Patra, "Sistem Analisa Harga Saham Menggunakan Algoritma Long Short Term Memory (LSTM)," *Sekolah Tinggi Manajemen Informatika & Komputer – LIKMI Bandung*, 2024. [Online]. Available: <https://library.likmi.ac.id/show/153/pdf>.
- I. H. Witten, E. Frank, and M. A. Hall, "Data Mining: Practical Machine Learning Tools and Techniques," Morgan Kaufmann, 2011. [Online]. Available: <https://www.elsevier.com/books/data-mining/witten/978-0-12-374856-0>
- K. A. Widiarto and E. W. Pamungkas, "Implementasi Algoritma LSTM untuk Prediksi Harga Saham pada Situs Yahoo Finance," *Universitas Muhammadiyah Surakarta*, 2024. [Online]. Available: <https://eprints.ums.ac.id/124206/1/Naskah%20Publikasi.pdf>.
- K. Kwanda et al., "Perbandingan LSTM dan Bidirectional LSTM pada Sistem Prediksi Harga Saham Berbasis Website," *Universitas Tarumanagara*, 2024. [Online]. Available: <https://jurnal.ranahresearch.com/index.php/R2/article/download/1255/1045/>.

El-Mujtama: Jurnal Pengabdian Masyarakat

Vol 5 No 3 (2025) 486 – 492 P-ISSN 2746-9794 E-ISSN 2747-2736

DOI: 10.47467/elmujtama.v5i3.7379

- K. Mauludin, "Model Prediksi Kecepatan Kurir dalam Pengiriman Surat Kabar," Eprints Unpak, 2023. [Online]. Tersedia: <https://eprints.unpak.ac.id/8115/1/Laporan%20-%20SKRIPSI%20-%2006512003%20-%20Kriti%20Mauludin.pdf>
- N. P. N. Kusuma et al., "Prediksi Harga Saham Blue Chip Menggunakan Algoritma Long Short-Term Memory (LSTM)," *Universitas Primakara*, 2024. [Online]. Available: <https://e-journalfb.ukdw.ac.id/index.php/jrak/article/download/6/6/31>.