



Tinjauan Algoritma Machine Learning untuk Prediksi Harga Saham: Studi Literatur

Miftahul Jannah

Teknik Informatika, Universitas Asahan, Indonesia

Corresponding Author

Nama Penulis: Miftahul Jannah

E-mail: miftahhstudy@gmail.com

Abstrak

Kemampuan untuk memprediksi harga saham secara akurat dari waktu ke waktu dapat memberikan keuntungan kompetitif yang signifikan dan mengurangi risiko investasi. Oleh karena itu, keuntungan diprediksi menggunakan prediksi yang akurat berdasarkan data yang valid. Metode penelitian yang digunakan dalam studi ini adalah tinjauan literatur yang komprehensif dan analisis peringkat. Hasil yang diharapkan dari studi ini adalah memberikan informasi tentang efektivitas relatif algoritma pembelajaran mesin dalam perdagangan mata uang, yang dapat berguna bagi investor dan pemilik bisnis. Orang-orang yang akan mendapatkan manfaat dari temuan studi ini adalah para investor dan pembuat kebijakan di sektor keuangan. Hasil yang diperoleh adalah MAPE atau RMSE dari algoritma LTSM dan Linear Regression, yang nilainya lebih kecil atau menunjukkan tingkat akurasi yang tinggi, sedangkan KNN Regression memberikan hasil yang lebih akurat.

Kata kunci - Machine Learning, Linear Regression, LTSM, KNN, Saham

Abstract

The ability to accurately predict stock prices over time can provide significant competitive advantages and reduce investment risks. Therefore, profits are predicted using accurate predictions based on valid data. The research methods used in this study are a comprehensive literature review and ranking analysis. The expected outcome of this study is to provide information on the relative effectiveness of machine learning algorithms in currency trading, which can be useful for investors and business owners. The people who will benefit from the findings of this study are investors and policymakers in the financial sector. The results obtained are MAPE or RMSE from the LSTM and Linear Regression algorithms, which have smaller values or show a high level of accuracy, while KNN Regression provides more accurate results.

Keywords - Machine Learning, Regresión Lineal, LSTM, KNN, Acciones

PENDAHULUAN

Pertumbuhan ekonomi suatu negara bukanlah hasil langsung dari volatilitas pasar modal. Pasar modal itu sendiri adalah sebuah organisasi dengan dua fungsi: ekonomi dan investasi. Pasar saham adalah tempat di mana barang diperjualbelikan dalam bentuk saham. Saham merupakan instrumen yang diakui secara luas dalam pasar modal. Dengan bantuan layanan ini, masyarakat umum dapat berinvestasi dalam saham perusahaan. Pasar modal digunakan untuk melacak perkembangan perdagangan saham baik dalam jangka panjang maupun pendek.

Saham adalah indikator kesehatan perusahaan yang berbentuk surat berharga. Investasi saham memiliki profil keuntungan tinggi dan risiko tinggi, di mana selain menawarkan potensi keuntungan besar, juga memiliki potensi risiko yang besar. Risiko tinggi yang muncul selama investasi menjadi peringatan bagi investor untuk melakukan analisis sebelum membeli saham. Salah satu komponen terpenting dalam ekonomi global yang memiliki dampak signifikan adalah pasar saham. Memprediksi harga saham telah menjadi alat penting bagi investor, analis keuangan, dan peneliti. Kemampuan untuk memprediksi harga saham secara akurat dari waktu ke waktu dapat memberikan keuntungan kompetitif yang signifikan dan mengurangi risiko investasi.

Seiring dengan kemajuan teknologi di era modern ini, peramalan nilai saham menarik perhatian para ekonom. Dengan kemampuan untuk memprediksi atau meramalkan harga saham, sebuah bisnis dapat lebih mudah menganalisis dan merekomendasikan langkah terbaik untuk menciptakan keputusan yang tepat dalam membeli atau menjual saham. Oleh karena itu, prediksi menjadi alat yang sangat penting untuk penelitian yang lebih efektif dan efisien. Mencatat riwayat harga saham harian, mingguan, bulanan, dan bahkan tahunan adalah cara prediksi dilakukan.

Ada beberapa faktor yang mempengaruhi harga saham, termasuk berita tentang politik, kondisi ekonomi nasional, inflasi, kondisi keuangan perusahaan, dan lainnya. Namun, jika kita kembali ke faktor dasar, penyebab utama volatilitas pasar saham adalah adanya perubahan di pasar dan permintaan informasi dalam jangka waktu tertentu. Ketika jumlah penawaran untuk suatu saham lebih besar daripada permintaannya, harga saham akan turun. Sebaliknya, ketika jumlah permintaan lebih besar daripada penawarannya, harga saham akan naik.

Salah satu metode yang dapat digunakan untuk memprediksi pergerakan harga saham dengan akurat adalah pembelajaran mesin (*machine learning*). Pembelajaran mesin adalah teknik yang memungkinkan komputer atau program memiliki kemampuan tertentu dengan menggunakan data, seperti teks, gambar, video, audio, atau angka, sebagai bahan pembelajaran untuk mengekstraksi pengetahuan dari data tersebut dan membantu manusia dalam menyelesaikan masalah atau memikirkan solusi.

Seiring dengan kemajuan teknologi dalam pembelajaran mesin, terdapat berbagai teknik dan model yang dapat digunakan untuk menganalisis kumpulan data tertentu. Untuk menentukan model pembelajaran mesin terbaik, analisis harus dilakukan berdasarkan data yang akan dianalisis. Beragam teknik analisis dan prediksi, termasuk penerapan teknik pembelajaran mesin, telah digunakan untuk membantu investor dalam pengambilan keputusan. Teknik ini terbukti efektif dalam menganalisis sejumlah besar data historis dan faktor pasar untuk memprediksi perubahan harga saham dari waktu ke waktu.

Kontribusi utama dari penelitian ini adalah memberikan analisis sistematis terhadap model-model saat ini serta implementasi model ensemble untuk analisis harga saham. Metode penelitian yang digunakan meliputi tinjauan literatur sistematis dan implementasi model ensemble.

Tujuan utama penelitian ini adalah menyediakan informasi komprehensif tentang artikel-artikel yang menggunakan pembelajaran mesin untuk menganalisis pergerakan pasar saham. Metode penelitian yang digunakan dalam studi ini mencakup tinjauan literatur yang komprehensif dan analisis peringkat. Hasil yang diharapkan dari penelitian ini adalah menyediakan informasi tentang efektivitas relatif algoritma pembelajaran mesin dalam perdagangan mata uang, yang dapat berguna bagi para

investor dan pemilik bisnis. Orang-orang yang akan mendapatkan manfaat dari temuan penelitian ini adalah investor dan pembuat kebijakan di sektor keuangan.

TINJAUAN PUSTAKA

Long Short-Term Memory (LSTM)

Komponen dari paradigma *Recurrent Neural Network (RNN)* adalah *Long Short-Term Memory (LSTM)*. Salah satu jenis jaringan saraf tiruan (ANN) yang didasarkan pada jaringan saraf *feedforward* disebut RNN. Jaringan saraf rekuren ini menggunakan pendekatan sistematis untuk menganalisis data input yang bersifat urutan (*sequential*). Meskipun RNN memberikan kinerja yang lebih baik dibandingkan dengan ANN tradisional, ia memiliki kekurangan yaitu tidak efektif pada data panjang karena masalah *vanishing gradient*. Untuk mengatasi masalah *vanishing gradient*, LSTM menggunakan sel memori yang terdiri dari lapisan input, tersembunyi, status sel, dan output.

KNN, atau K-Nearest Neighbors

KNN adalah metode yang sederhana namun efektif untuk klasifikasi dan regresi. Metode ini mencari k titik data terdekat dari titik data yang perlu diprediksi, kemudian membuat prediksi berdasarkan nilai dari titik-titik terdekat tersebut. KNN dapat digunakan untuk memprediksi harga saham dengan mencari saham yang memiliki data harga historis yang serupa dengan saham yang perlu diprediksi.

Algoritma Regresi Linear

Salah satu metode penelitian yang menggunakan hubungan lurus untuk menggambarkan hubungan antara dua variabel atau lebih disebut dengan regresi linear. Regresi linear dengan satu variabel bebas disebut regresi linear sederhana, sedangkan regresi linear dengan lebih dari satu variabel bebas disebut regresi linear berganda.

RMSE

RMSE adalah metode yang digunakan untuk melakukan pengukuran dengan meminimalkan perbedaan antara nilai yang diamati yang dihasilkan dari prediksi model dalam memperkirakan nilai yang diamati. Keakuratan dalam memperkirakan nilai tertentu ditunjukkan oleh hasil RMSE yang kecil. Dapat dikatakan bahwa nilai RMSE yang lebih kecil untuk suatu estimasi lebih akurat jika dibandingkan dengan nilai RMSE yang lebih besar. Dalam analisis data menggunakan teknik pembelajaran mesin seperti ini, parameter RMSE sangat penting untuk menentukan akurasi dari model yang diberikan.

METODE

Metodologi penelitian ini didasarkan pada studi literatur, dengan fokus pada data atau penelitian sebelumnya di mana temuan-temuannya digunakan sebagai panduan untuk menentukan algoritma pembelajaran mesin terbaik yang dapat digunakan dalam memprediksi harga saham.

Tinjauan literatur berpusat pada penggunaan algoritma pembelajaran mesin dalam memprediksi harga saham. Artikel ini, yang diterbitkan dalam rentang tahun ketiga 2022–2024, bertujuan untuk menyoroti kemajuan teknologi. Sumber utama penelitian literatur adalah Google Scholar, yang menyediakan akses luas ke artikel ilmiah, jurnal, dan makalah penelitian di berbagai bidang.

Ada tiga algoritma yang menjadi fokus analisis: *K-Nearest Neighbors Regression*, Algoritma Linear Regression, dan Algoritma *Long Short-Term Memory (LSTM)*.

Prosedur Seleksi

Tahap 1: Penelitian awal dilakukan berdasarkan abstrak dan judul untuk menilai relevansi fokus utama penelitian, yaitu prediksi harga saham.

Tahap 2: Artikel dievaluasi secara menyeluruh dalam hal metodologi, data, dan hasil analisis.

Tahap 3: Literatur dikategorikan berdasarkan algoritma pembelajaran mesin yang digunakan, yaitu *K-Nearest Neighbors Regression*, Algoritma *Linear Regression*, dan Algoritma *Long Short-Term Memory (LSTM)*.

Indikator Penilaian

1. Analisis Berdasarkan *Mean Absolute Percentage Error (MAPE)* dan *Root Mean Square Error (RMSE)*: Digunakan untuk mengevaluasi tingkat akurasi dari masing-masing algoritma.
2. Ukuran atau Kebutuhan Data pada Implementasi: Karena setiap algoritma memiliki kemampuan unik yang sesuai dengan kasus tertentu, kebutuhan data menjadi aspek penting dalam menentukan kecocokan algoritma.

Berdasarkan indikator ini, kondisi saat ini akan dievaluasi berdasarkan penggunaan algoritma. Evaluasi ini bertujuan untuk mendukung pemilihan algoritma yang optimal sesuai dengan kondisi yang mendekati performa terbaiknya (kinerjanya).

PEMBAHASAN

Terdapat tiga penelitian yang menjadi sumber studi literatur:

1. Judul: *Penerapan Machine Learning Dalam Melakukan Prediksi Harga Saham Pt. Bank Mandiri (Persero) Tbk Dengan Algoritma Linear Regression*

Data: 1255 harga saham historis dari PT. Bank Mandiri Tbk, mencakup periode 5 tahun, dari 1 November 2018 hingga 29 November 2023.

Hasil: Berdasarkan hasil vektor performa dari model prediksi saham *RapidMiner* yang menggunakan algoritma regresi linear, hasil yang diperoleh menunjukkan performa yang baik:

RMSE: 0.005 +/- 0.000

Kesalahan absolut: 0.004 +/- 0.003

Kesalahan relatif: 2.04% +/- 3.49%

Korelasi: 0.999

Korelasi kuadrat: 0.998

2. Judul: *Implementasi Algoritma K-Nearest Neighbors Regression Dalam Memprediksi Harga Saham*

Data: 1547 data saham PT Bank Negara Indonesia Tbk dengan nama BBNI.JK.

Hasil: RMSE menunjukkan hasil tertinggi pada rasio 70:30. Berdasarkan hasil ini, data dibandingkan dengan rasio 70:30 dan RMSE tercatat paling rendah untuk semua variabel indikator. Model hasil menggunakan regresi KNN dengan tiga variabel indikator yang berbeda menunjukkan hasil terbaik, dengan RMSE sebesar 169 dan akurasi 98,8%. Hasil prediksi menggunakan satu indikator menunjukkan hasil terbaik. Prediksi menggunakan KNN memiliki rentang antara 40 hingga 100 untuk setiap variabel indikator.

Tabel 1.

Tabel Hasil Pemodelan KNN

	1 Indikator	2 Indikator	3 Indikator	4 Indikator
K – Optimal	7	6	6	6
RMSE	168	177	177	177
Akurasi	99.17%	99.14%	99.14%	99.07%

Sumber : Data Artikel Penelitian Terkait

3. Judul: *Prediksi Harga Saham Syariah Menggunakan Algoritma Long Short-Term Memory (LSTM)*

Data: Data harga saham lima emiten syariah yang meliputi PT Aneka Tambang Tbk (ANTM), Erajaya Swasembada Tbk (ERAA), Kalbe Farma Tbk (KLBF), Semen Indonesia (Persero) Tbk (SMGR), dan Wijaya Karya (Persero) Tbk (WIKA). Semua data yang digunakan mencakup beberapa tahun, dimulai dari Desember 2016 hingga Desember 2021.

Prosedur: Uji coba dilakukan dengan menggunakan beberapa parameter pada *layer*, *epoch*, dan *time step* untuk memperoleh model prediksi terbaik atau paling akurat. Setiap emiten memiliki data yang sedikit berbeda, sehingga setiap emiten akan memiliki model terbaik dengan pengaturan *layer*, *epoch*, dan *time step* yang berbeda.

Hasil:

This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license



PT Aneka Tambang Tbk (ANTM): MAPE sebesar 2.64 berdasarkan hasil studi menggunakan metode *Long Short-Term Memory (LSTM)*.

Erajaya Swasembada Tbk (ERAA): MAPE sebesar 2.24.

Kalbe Farma Tbk (KLBF): MAPE sebesar 1.51.

Semen Indonesia (Persero) Tbk (SMGR): MAPE sebesar 1.83.

Wijaya Karya (Persero) Tbk (WIKA): MAPE sebesar 2.66.

Hasil-hasil ini menunjukkan variasi tingkat akurasi dalam prediksi harga saham masing-masing emiten menggunakan algoritma LSTM.

KESIMPULAN

Berdasarkan hasil studi literatur, dapat diamati perbedaan dalam algoritma pembelajaran mesin. Kita dapat melihat bahwa MAPE atau RMSE dari algoritma LSTM dan Regresi Linear lebih kecil atau menunjukkan akurasi yang tinggi, sementara Regresi KNN lebih akurat. Membandingkan kompleksitas data juga menunjukkan bahwa meskipun data yang digunakan dalam LSTM lebih besar dan lebih lengkap, Regresi Linear memiliki data yang lebih kecil tetapi akurasi yang lebih tinggi; ini bisa menjadi alternatif ketika menggunakan algoritma pembelajaran mesin untuk memprediksi harga saham dengan data yang lebih kecil. Sebaliknya, Regresi KNN direkomendasikan untuk dataset dengan jumlah indikator yang lebih sedikit, karena meskipun jumlah data lebih banyak dibandingkan dengan penggunaan Regresi Linear pada artikel pertama, nilai RMSE lebih tinggi. Kita dapat mengamati perbedaan antara penggunaan berbagai algoritma pembelajaran mesin, di mana algoritma akan tampil terbaik jika diterapkan sesuai dengan kemampuan dan kemampuannya untuk menangani situasi atau kasus yang diberikan. Untuk penelitian lebih lanjut, perbandingan dapat dilakukan di area lain, seperti menentukan harga barang.

DAFTAR PUSTAKA

- Afrianto, N., Fudholi, D. H., & Rani, S. (2021). Prediksi harga saham menggunakan BiLSTM dengan faktor sentimen publik. *JURNAL RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6(1), 41–46.
- Budiprasetyo, G., Hani'ah, M., & Aflah, D. Z. (2022). Prediksi harga saham syariah menggunakan algoritma Long Short-Term Memory (LSTM). *Jurnal Nasional Teknologi dan Sistem Informasi*, 8(3), 164–172.
- Eka, P. (2020). Implementasi Support Vector Machine pada prediksi harga saham gabungan (IHSG). *Jurnal Ilmiah Teknologi dan Rekayasa*, 25(1), 24–38.
- Hastomo, W., Sutarno, & Sudjiran. (2022). Analisis risiko investasi dan prediksi saham menggunakan algoritme machine learning. *Jurnal Ilmiah KOMPUTASI*, 21(3), 453–462.
- Kuswanto, A. D., Andro, R. H., Januar, G. A., Sandrio, R. O., Andera, N., & Cahya, T. D. (2024). Estimasi harga saham menggunakan algoritma regresi linear. *JURNAL INFORMATIKA*, 11(1), 1–8.
- Mulyadi, M. (2011). Penelitian kuantitatif dan kualitatif serta pemikiran dasar menggabungkannya. *Jurnal Studi Komunikasi dan Media*, 15(1), 127–138.
- Nurfauzan, D., & Fatimah, T. (2022). Implementasi algoritma K-Nearest Neighbors Regression dalam memprediksi harga saham. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 576–584.
- Rozi, F., & Junianto, M. B. S. (2023). Penerapan machine learning untuk prediksi harga saham PT. Telekomunikasi Indonesia Tbk menggunakan algoritma K-Nearest Neighbors. *Jurnal Informatika MULTI*, 1(1), 18–24.
- Sapanji, R. A. E., Targa, V., Lestari, S., Murnawan, & Samihardjo, R. (2023). Prediksi indeks Bursa Efek Indonesia 2023 pendekatan ARIMA, machine learning dengan R programming. *Jurnal Manajemen Informatika (JAMIKA)*, 13(2), 163–177.

- Santoso, R. S., & Dewi, F. K. S. (2024). Konfigurasi model Prophet untuk prediksi harga saham sektor teknologi di Indonesia yang akurat. *Jurnal Buana Informatika*, 15(1), 50–58.
- Sari, E. P., Bachri, M. S., Atnang, M., Sahriani, & Fajar, N. (2024). Studi literatur deep learning dan machine learning untuk analisis dan prediksi pasar saham: Metodologi, representasi data, dan studi kasus. *Jurnal Teknologi dan Sains Modern*, 1(1), 19–28.
- Sudriyanto, F. S., & Fitriani, N. (2023). Perbandingan performa model machine learning Support Vector Machine, Neural Network, dan K-Nearest Neighbors dalam prediksi harga saham. *Jurnal Advance Research Informatika*, 2(1), 13–21.
- Syahfutri, Y., Munthe, I. R., & Harahap, S. Z. (2023). Analisis machine learning algoritma regresi linear untuk memprediksi saham di Bank BRI di Bursa Saham Indonesia. *Jurnal TEKINKOM*, 6(1), 81–87.
- Thabibi, A., & Supriyanto, R. (2023). Perbandingan model multiple linear regression dan decision tree regression (Studi kasus: Prediksi harga saham Telkom, Indosat, dan XL). *Jurnal Ilmiah Teknologi dan Rekayasa*, 28(1), 78–92.
- Triya, P., Suarna, N., & Nuris, N. D. (2024). Penerapan machine learning dalam melakukan prediksi harga saham PT. Bank Mandiri (Persero) Tbk dengan algoritma linear regression. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 1207–1214.
- Utomo, W. C. (2023). Prediksi pergerakan saham BBRI di tengah isu ancaman resesi 2023 dengan pendekatan machine learning. *Jurnal Teknologi dan Manajemen Informatika*, 9(1), 20–27.
- Waluyo, D. E., Kinasih, H. W., Paramita, C., Pergiwati, D., Rafrastara, F. A., & Nohan, R. (2024). Implementasi algoritma regresi pada machine learning untuk prediksi indeks harga saham gabungan. *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, 9(1), 12–17.
- Wiseto, P. A. (2021). Optimasi parameter input pada artificial neural network untuk meningkatkan akurasi prediksi indeks harga saham. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 10(2), 211–216.