

Studi Penerapan Deep Learning untuk Prediksi Pergerakan Indeks Harga Saham Gabungan

Irene Karijadi*, Yoseph Widjaja, Dian Retno Sari Dewi, Luh Juni Asrini

Universitas Katolik Widya Mandala Surabaya, Indonesia

Email: irenekarijadi92@gmail.com*

Keywords	Abstrak
LSTM, IHSG, Prediksi, Univariate, Multivariate.	Penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model prediksi pergerakan Indeks Harga Saham Gabungan (IHSG) menggunakan metode Long Short-Term Memory (LSTM) sebagai salah satu pendekatan deep learning yang efektif dalam menangkap pola deret waktu nonlinier. Penelitian ini menguji tiga pendekatan model, yaitu model Univariate yang hanya menggunakan data historis IHSG, model Multivariate All Feature yang mengintegrasikan seluruh variabel eksternal, serta model Multivariate Selected Feature yang menggunakan variabel eksternal terpilih. Variabel eksternal yang dipertimbangkan meliputi harga emas dunia, harga minyak dunia, nilai tukar rupiah terhadap dolar Amerika Serikat, serta indeks saham internasional yang berpengaruh terhadap pasar modal Indonesia. Evaluasi kinerja model dilakukan menggunakan indikator Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), dan Mean Absolute Error (MAE). Hasil penelitian menunjukkan bahwa model Multivariate All Feature memberikan performa terbaik dengan nilai MAPE sebesar 0,76, RMSE sebesar 66,72, dan MAE sebesar 51,58, yang secara konsisten lebih rendah dibandingkan dua model lainnya. Uji signifikansi menggunakan ANOVA dan Tukey HSD mengonfirmasi adanya perbedaan kinerja yang signifikan antar model. Temuan ini mengindikasikan bahwa integrasi variabel eksternal mampu meningkatkan akurasi prediksi IHSG secara signifikan. Penelitian ini diharapkan dapat menjadi referensi bagi investor, analis pasar modal, dan peneliti dalam pengembangan sistem pendukung keputusan berbasis kecerdasan buatan di pasar saham Indonesia.
Keyword	Abstract
LSTM, JCI, Prediction, Univariate, Multivariate	<i>This study aims to develop and evaluate a prediction model for the movement of the Composite Stock Price Index (JCI) using the Long Short-Term Memory (LSTM) method as one of the effective deep learning approaches in capturing nonlinear time series patterns. This study tested three model approaches, namely the Univariate model which only uses historical JCI data, the Multivariate All Feature model which integrates all external variables, and the Multivariate Selected Feature model which uses selected external variables. External variables considered include world gold prices, world oil prices, rupiah exchange rates against the United States dollar, and international stock indices that affect the Indonesian capital market. The model performance evaluation was carried out using the Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) indicators. The results showed that the Multivariate All Feature model provided the best performance with a MAPE value of 0.76, an RMSE of 66.72, and an MAE of 51.58, which was consistently lower than the other two models. Significance tests using ANOVA and Tukey HSD confirmed significant performance differences between models. These findings indicate that the integration of</i>

external variables is able to significantly increase the accuracy of JCI predictions. This research is expected to be a reference for investors, capital market analysts, and researchers in the development of artificial intelligence-based decision support systems in the Indonesian stock market.



PENDAHULUAN

Di era globalisasi saat ini, pasar modal menjadi salah satu indikator penting dalam perekonomian suatu negara. Indeks Harga Saham Gabungan (IHSG), yang mencerminkan pergerakan harga saham-saham yang terdaftar di Bursa Efek Indonesia (BEI), menjadi salah satu acuan utama bagi investor dalam menilai kinerja pasar saham Indonesia. Sebagai indikator ekonomi, IHSG memiliki peranan yang sangat penting dalam menggambarkan kondisi pasar modal secara keseluruhan. Fluktuasi pergerakan IHSG dapat memberikan gambaran tentang arah pertumbuhan ekonomi negara. Seiring dengan itu, prediksi terhadap pergerakan IHSG menjadi sangat penting karena memberikan informasi bagi investor, pedagang, dan lembaga keuangan untuk mengambil keputusan yang tepat, mengelola risiko, serta mengoptimalkan strategi investasi yang dapat meningkatkan keuntungan (Madi, 2018).

Salah satu cara untuk memprediksi pergerakan IHSG di masa depan adalah dengan menggunakan teknik prediksi harga yang memperhitungkan riwayat pergerakan IHSG sebelumnya. Pendekatan ini umum digunakan karena data pergerakan IHSG merupakan data time series yang mencerminkan pola historis pergerakan pasar. Namun demikian, pergerakan IHSG tidak hanya dipengaruhi oleh riwayat pergerakannya sendiri, tetapi juga oleh berbagai faktor eksternal seperti kondisi ekonomi makro dan mikro. Faktor-faktor ini menyebabkan pergerakan IHSG menjadi sangat kompleks, dinamis, serta bersifat nonlinier dan nonstasioner (Hassan, 2019). Oleh karena itu, memprediksi pergerakan IHSG secara akurat merupakan tantangan tersendiri yang membutuhkan pendekatan dan model yang tepat. Dengan adanya prediksi terhadap IHSG, investor diharapkan dapat memperoleh informasi tambahan yang bermanfaat sebagai pertimbangan dalam merumuskan strategi investasi dan mengelola risiko secara lebih terukur.

Data pergerakan IHSG merupakan data time series, sehingga prediksi pergerakan IHSG dapat dilakukan dengan menggunakan data time series pada hari-hari sebelumnya untuk memprediksi pergerakan IHSG pada masa depan (Idrees, Alam, dan Agarwal, 2019). Ada berbagai metode yang dapat mengolah data time series dan digunakan untuk memprediksi pergerakan IHSG, mulai dari pendekatan statistik tradisional hingga teknik machine learning yang lebih baru. Salah satu metode yang paling populer adalah Autoregressive Integrated Moving Average (ARIMA), yang digunakan untuk menganalisis dan memprediksi data time series berdasarkan pola masa lalu harga saham (Box dan Jenkins, 1970). Selain ARIMA, model Long Short-Term Memory (LSTM), yang merupakan jenis jaringan saraf tiruan (neural network) dalam deep learning, semakin banyak digunakan untuk menangkap hubungan jangka panjang dalam data time series dan dapat mengatasi ketergantungan temporal yang kompleks dalam harga saham (Hochreiter dan Schmidhuber, 1997). Metode lain yang sering digunakan adalah Support Vector Machines (SVM), yang efektif dalam menemukan pola non-linear dalam data harga saham, serta Random Forests, yang dapat menggabungkan berbagai model prediksi untuk menghasilkan hasil yang lebih akurat (Krauss, Do, dan Huck, 2017).

Dari berbagai macam metode yang ada untuk memprediksi data time series, salah satu pendekatan yang populer adalah penggunaan LSTM. LSTM adalah jenis jaringan saraf tiruan yang termasuk dalam kategori Recurrent Neural Network (RNN), dirancang untuk mengatasi masalah dalam memproses data sekuensial yang memiliki ketergantungan temporal, seperti data time series. LSTM mampu menangkap pola jangka panjang dalam data dengan cara mengingat informasi yang relevan melalui gerbang-gerbang khusus, yang membedakannya dari RNN tradisional yang sering mengalami vanishing gradient problem (Hochreiter dan Schmidhuber, 1997). Metode LSTM dipilih dalam penelitian ini karena kemampuannya untuk mengatasi ketergantungan jangka panjang dan non-linier dalam pergerakan IHSG yang sering dipengaruhi oleh berbagai faktor eksternal. Berbeda dengan metode tradisional seperti ARIMA yang membutuhkan asumsi stasioneritas, LSTM dapat bekerja langsung dengan data mentah dan memberikan prediksi yang lebih akurat dan relevan, sesuai dengan dinamika pasar saham yang sangat kompleks dan dinamis.

Penelitian sebelumnya telah banyak menggunakan metode LSTM untuk memprediksi data time series, sebagai contoh, penelitian oleh Siami-Namini, Tavakoli, dan Siami Namin (2019) menunjukkan bahwa LSTM dapat mengurangi error prediksi hingga 84%–87% lebih baik dibandingkan dengan metode ARIMA, yang cenderung terbatas dalam menangani ketergantungan jangka panjang dan pola non-linier dalam data time series. Sementara itu, penelitian oleh Gao, Chai, dan Liu (2018) juga membuktikan bahwa LSTM mengungguli model-model lain seperti *Moving Average* (MA) dan *Support Vector Machine* (SVM) dalam memprediksi pergerakan S&P500. Meskipun penelitian-penelitian tersebut memberikan bukti bahwa LSTM efektif dalam prediksi data time series, penelitian ini akan melangkah lebih jauh dengan membandingkan ketiga pendekatan *Univariate*, *Multivariate All Feature*, dan *Multivariate Selected Feature*.

Dalam penelitian ini, selain menggunakan data pergerakan IHSG itu sendiri, faktor eksternal yang memengaruhi pergerakan IHSG, seperti Harga Emas Dunia (Gumilang, 2014), Harga Minyak Dunia (Gumilang, 2014), Nilai Tukar Rupiah terhadap Dolar Amerika Serikat (Gumilang, 2014), Indeks Saham Hong Kong (Chabachib dan Witjaksono, 2011), Indeks Saham Jepang (Chabachib dan Witjaksono, 2011), Indeks Saham Amerika (Putra dan Nurmatias, 2024), dan Indeks Morgan Stanley Capital Internasional (MSCI) Emerging Markets (Invesnow, 2024). akan dimasukkan dalam model Multivariate untuk melihat apakah kombinasi berbagai variabel eksternal dapat meningkatkan akurasi prediksi pergerakan IHSG dibandingkan dengan menggunakan data historis pergerakan IHSG saja.

Penelitian ini bertujuan untuk mengembangkan model prediksi pergerakan IHSG dengan tiga pendekatan berbeda. Fokus pertama adalah mengembangkan model prediksi pergerakan IHSG secara LSTM Univariate menggunakan data historis pergerakan IHSG (^JKSE) itu sendiri. Kedua, penelitian ini juga akan mengeksplorasi penggunaan pendekatan Multivariate all feature dengan memperkenalkan faktor eksternal seperti Harga Emas Dunia (GC=F), Harga Minyak Dunia (CL=F), Nilai Tukar Rupiah terhadap Dolar Amerika Serikat (IDR=X), Indeks Saham Hong Kong (^HSI), Indeks Saham Jepang (^N225), Indeks Saham Amerika (^DJI), dan Indeks *MSCI Emerging Markets* (EEM) yang diperkirakan dapat memberikan wawasan lebih mendalam mengenai pergerakan IHSG. Ketiga, penelitian ini juga akan mengembangkan model prediksi pergerakan IHSG secara Multivariate Selected Feature menggunakan metode LSTM dengan mempertimbangkan faktor eksternal yang dipilih

menggunakan teknik feature selection. Keempat, dengan membandingkan ketiga pendekatan ini, penelitian ini bertujuan untuk melihat efektivitas model Univariate, Multivariate All Feature, dan Multivariate Selected Feature dalam memberikan akurasi prediksi pergerakan IHSG yang lebih akurat dan dapat diandalkan.

METODE PENELITIAN

Prediksi pergerakan IHSG menjadi tantangan besar karena fluktuasi pasar yang dipengaruhi oleh berbagai faktor eksternal dan internal yang kompleks. Metode tradisional seperti ARIMA terbukti kurang efektif dalam menangani ketergantungan jangka panjang dan sifat non-linier dari data time series IHSG. Oleh karena itu, model LSTM dipilih untuk memprediksi pergerakan IHSG karena kemampuannya dalam mengatasi dinamika pasar yang kompleks. Namun, sebagian besar penelitian sebelumnya hanya menggunakan pendekatan LSTM Univariate, yang hanya mempertimbangkan data pergerakan IHSG itu sendiri. Permasalahan utama dalam penelitian ini adalah bagaimana mengembangkan model prediksi yang lebih akurat dengan pendekatan LSTM Multivariate All Feature yang melibatkan faktor eksternal, seperti Harga Emas Dunia, Harga Minyak Dunia, Nilai Tukar Rupiah terhadap Dolar Amerika Serikat, dan Indeks pasar global. Lebih lanjut, penelitian ini juga akan mengeksplorasi pendekatan LSTM Multivariate Selected Feature, yang mana feature selection digunakan untuk memilih variabel eksternal yang paling relevan, guna meningkatkan akurasi pergerakan IHSG dibandingkan dengan hanya menggunakan pendekatan Univariate.

Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan dengan memanfaatkan platform Yahoo Finance, yang menyediakan data historis pasar saham secara komprehensif. Data yang dikumpulkan mencakup harga penutupan harian IHSG ($^{\wedge}JKSE$), serta data dari beberapa variabel prediktor lainnya, yaitu Harga Emas Dunia ($GC=F$), Harga Minyak Dunia ($CL=F$), Nilai Tukar Rupiah terhadap Dolar AS ($IDR=X$), Indeks Saham Hong Kong ($^{\wedge}HSI$), Indeks Saham Jepang (N225), Indeks Saham Amerika ($^{\wedge}DJI$), dan Indeks MSCI Emerging Markets (EEM). Rentang waktu data yang diambil adalah dari 1 Januari 2005 hingga 1 Januari 2025, karena untuk melakukan prediksi menggunakan model LSTM, berdasarkan studi yang dilakukan oleh Sezer, Gudelek, dan Ozbayoglu (2020), 3 hingga 5 tahun data harian sudah cukup untuk menghasilkan model yang dapat belajar pola pergerakan harga dengan baik. Data yang lebih lama memungkinkan model menangkap siklus pasar yang lebih panjang, namun terlalu banyak data juga dapat menyebabkan overfitting jika pola-pola lama tidak lagi relevan dengan kondisi pasar saat ini. LSTM bekerja dengan baik pada data deret waktu dan lebih efektif dengan jumlah data yang lebih banyak untuk memahami hubungan jangka panjang. Sebagai referensi, Sezer, dkk. (2020) menemukan bahwa untuk prediksi harga saham, penggunaan data harian selama 3-5 tahun terakhir sudah memadai untuk model LSTM yang akurat. Oleh karena itu, disarankan untuk menggunakan minimal 3 tahun data harian untuk mendapatkan hasil yang optimal dalam prediksi harga saham (Sezer, dkk., 2020).

Data diunduh menggunakan library yfinance dalam Python, yang memudahkan proses pengumpulan data dari Yahoo Finance. Pemilihan variabel prediktor didasarkan pada bukti empiris dan teori yang menunjukkan pengaruh signifikan terhadap pergerakan IHSG:

Harga Emas Dunia (GC=F) dipilih karena emas sering dianggap sebagai aset safe haven, yang menjadi pilihan investasi saat terjadi ketidakpastian ekonomi. Harga emas memiliki hubungan negatif dengan pasar saham: saat pasar saham mengalami penurunan, investor cenderung beralih ke emas. Oleh karena itu, fluktuasi harga emas dapat mempengaruhi pergerakan IHSG, terutama dalam situasi ketidakpastian ekonomi global atau krisis finansial (Gumilang, 2014).

Harga Minyak Dunia (CL=F) dipilih karena Indonesia adalah negara yang sangat bergantung pada impor energi, khususnya minyak. Perubahan harga minyak mempengaruhi biaya produksi, inflasi, dan daya beli masyarakat, yang pada gilirannya dapat mempengaruhi kinerja saham di pasar Indonesia. Kenaikan harga minyak biasanya akan menekan margin keuntungan perusahaan-perusahaan yang bergantung pada energi, menyebabkan penurunan harga saham, yang tercermin dalam pergerakan IHSG (Putra dan Nurmatias, 2024).

Nilai Tukar Rupiah terhadap Dolar AS (IDR=X) menjadi variabel penting karena fluktuasi nilai tukar dapat mempengaruhi arus modal asing yang masuk atau keluar dari pasar saham Indonesia. Kelemahan rupiah terhadap dolar biasanya akan meningkatkan biaya impor dan mengurangi daya beli domestik, yang dapat menekan kinerja pasar saham. Sebaliknya, penguatan rupiah dapat meningkatkan sentimen pasar dan memperbaiki kinerja saham. Oleh karena itu, nilai tukar rupiah sangat relevan dalam memprediksi pergerakan IHSG, terutama dalam konteks arus investasi asing (Putra dan Nurmatias, 2024).

Indeks Saham Hong Kong (^HSI) dipilih karena mencerminkan pasar saham di Hong Kong, yang merupakan salah satu ekonomi terbesar di Asia. Perubahan signifikan pada indeks ini sering kali diikuti oleh pergerakan yang serupa di pasar saham Indonesia, mengingat hubungan ekonomi yang erat antara Indonesia dan pasar Asia. Misalnya, penurunan signifikan pada Indeks Saham Hong Kong dapat mempengaruhi pasar Indonesia melalui aliran modal internasional dan sentimen investor (Gumilang, 2014).

Indeks Saham Jepang (^N225) dipilih karena mencerminkan pasar saham Jepang, yang merupakan salah satu ekonomi terbesar di dunia. Pergerakan Saham Jepang dapat mempengaruhi sentimen pasar saham global dan mempengaruhi aliran modal internasional, yang pada gilirannya dapat mempengaruhi pergerakan IHSG. Penurunan atau kenaikan signifikan pada Indeks Saham Jepang dapat menjadi indikator perubahan tren global yang berdampak pada pasar saham Indonesia (Chabachib dan Witjaksono, 2011).

Indeks Saham Amerika (^DJI) dipilih karena mencerminkan kondisi pasar saham di Amerika Serikat, yang memiliki pengaruh besar terhadap ekonomi global. Perubahan signifikan pada Indeks Saham Amerika sering kali menciptakan sentimen negatif atau positif yang mempengaruhi pasar saham Indonesia, melalui pengaruh dalam aliran modal dan kepercayaan investor internasional (Putra dan Nurmatias, 2024).

Indeks MSCI Emerging Markets (EEM) merupakan indeks yang mencakup saham-saham dari negara berkembang, termasuk Indonesia. Perubahan dalam bobot atau komposisi saham Indonesia dalam indeks ini dapat mempengaruhi aliran dana asing dan pergerakan Indeks Harga Saham Gabungan (IHSG). Ketika MSCI mengurangi bobot saham Indonesia dalam indeks EEM, seperti yang terjadi pada 2024 ketika bobot Indonesia turun dari 2% menjadi 1,5%, hal ini dapat menyebabkan penurunan aliran dana asing yang berinvestasi di pasar saham Indonesia, karena banyak investor global yang mengandalkan indeks EEM dalam keputusan investasi mereka. Penurunan ini berpotensi menekan IHSG, seperti yang tercatat

dalam reaksi pasar saat itu (Invesnow, 2024). Sebaliknya, jika bobot saham Indonesia dalam indeks EEM meningkat atau Indonesia dimasukkan dalam kategori yang lebih tinggi, hal ini dapat menarik lebih banyak aliran dana asing dan mendukung penguatan IHSG (Bareksa, 2024).

Preprocessing Data

Preprocessing data merupakan tahap kritis dalam penelitian ini, yang bertujuan untuk mempersiapkan data mentah agar sesuai untuk pelatihan model LSTM. Langkah pertama yang dilakukan adalah menangani nilai yang hilang (missing values) dengan menghapus baris yang mengandung nilai hilang menggunakan fungsi `dropna`). Selanjutnya, scaling data menggunakan `MinMaxScaler` untuk mengubah rentang nilai semua variabel menjadi antara 0 dan 1, agar model LSTM dapat memproses data secara efisien. Proses ini penting karena model LSTM lebih efektif ketika data berada dalam skala yang seragam. Setelah itu, scaling data dilakukan dengan membagi data menjadi dua bagian yaitu, 80% training dan 20% testing, dengan tujuan untuk melatih model pada data yang sudah ada dan menguji performanya pada data yang belum pernah dilihat sebelumnya.

Selain itu, teknik time lag digunakan untuk memanfaatkan informasi dari 28 hari sebelumnya untuk memprediksi pergerakan IHSG di masa depan. Penggunaan 28 time lag dipilih karena data yang digunakan merupakan data harian (daily), sehingga periode 28 hari mencakup hampir satu siklus bulanan yang dapat menangkap pola musiman dan tren pergerakan harga yang sering berulang setiap bulan (Siami-Namini, Tavakoli, dan Siami Namin, 2019). Dengan demikian, data siap untuk digunakan dalam pelatihan model, memastikan bahwa proses analisis dapat dilakukan dengan data yang bersih, terstandarisasi, dan relevan.

LSTM Univariate

Model LSTM univariate dalam penelitian ini dilatih secara khusus menggunakan data pergerakan IHSG (^JKSE) sebagai satu-satunya input. Model ini dirancang untuk mempelajari pola dan tren historis dari pergerakan IHSG, dengan tujuan memprediksi pergerakan IHSG di masa depan hanya berdasarkan riwayat pergerakannya sendiri. Untuk mengoptimalkan performa model, proses `GridSearch` dilakukan untuk menemukan kombinasi terbaik dari hyperparameter epoch dan batch size. `GridSearch` mengevaluasi berbagai kombinasi nilai epoch dan batch size yang telah ditentukan dalam ruang pencarian yang ditetapkan, dengan tujuan mengurangi kesalahan prediksi dan meningkatkan akurasi model. Dengan melakukan `GridSearch`, model LSTM ini diharapkan dapat memberikan hasil yang lebih akurat dalam memprediksi pergerakan IHSG berdasarkan data historisnya.

Uji Multikolinieritas

Uji multikolinieritas dilakukan pada seluruh variabel prediktor dalam penelitian ini, yaitu GC=F , CL=F , IDR=X , ^HSI , ^N225 , ^DJI , dan EEM untuk memastikan tidak ada korelasi tinggi antar variabel yang dapat mengganggu kestabilan model. Hasil uji multikolinieritas ini digunakan untuk dua hal utama. Pertama, untuk membangun model LSTM Multivariate All Feature yang menggunakan semua variabel prediktor yang telah disaring berdasarkan uji VIF. Kedua, hasil uji ini digunakan dalam tahap feature selection menggunakan

RFE, di mana variabel-variabel yang paling relevan dipilih untuk membangun model LSTM Multivariate Selected Feature. Dengan demikian, uji multikolinieritas memastikan bahwa model yang dibangun lebih efisien dan mengurangi risiko overfitting.

LSTM Multivariate All Feature

Model LSTM *Multivariate All Feature* dalam penelitian ini dilatih menggunakan data yang mencakup pergerakan IHSG dan variabel-variabel prediktor lainnya. Arsitektur model LSTM Multivariate All Feature serupa dengan model Univariate, tetapi disesuaikan dengan jumlah variabel input yang lebih banyak. Untuk mengoptimalkan performa model, proses GridSearch juga diterapkan pada model Multivariate All Feature ini. GridSearch digunakan untuk mencari kombinasi terbaik dari hyperparameter epoch dan batch size dengan menguji berbagai nilai yang telah ditentukan sebelumnya. Dengan menggunakan GridSearch, model dapat memilih kombinasi yang menghasilkan performa terbaik, mengurangi kesalahan prediksi, dan meningkatkan akurasi dalam memprediksi pergerakan IHSG.

Feature Selection

Pemilihan variabel dalam penelitian ini dilakukan menggunakan metode Recursive Feature Elimination (RFE). Metode ini secara iteratif menghilangkan variabel-variabel yang dianggap kurang penting, hingga hanya tersisa variabel-variabel yang paling relevan untuk memprediksi pergerakan IHSG.

LSTM Multivariate Selected Feature

Setelah dilakukan feature selection, model LSTM Multivariate Selected Feature dikembangkan dengan menggunakan variabel-variabel terpilih dari tahap sebelumnya. Dengan menggunakan variabel yang relevan, model ini diharapkan dapat memberikan prediksi pergerakan IHSG yang lebih tepat, karena hanya mempertimbangkan variabel-variabel yang signifikan dalam proses prediksi. Untuk mengoptimalkan performa model, GridSearch diterapkan untuk mencari kombinasi terbaik dari hyperparameter epoch dan batch size. Dengan menguji berbagai kombinasi nilai yang telah ditentukan sebelumnya, GridSearch membantu dalam memilih parameter yang menghasilkan performa terbaik, sehingga model dapat memberikan prediksi yang lebih akurat dengan menggunakan variabel yang paling relevan dan signifikan untuk pergerakan IHSG.

Evaluasi

Evaluasi model dalam penelitian ini dilakukan dengan membandingkan tiga metrik evaluasi, yaitu Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), dan Mean Absolute Error (MAE). MAPE dipilih karena memberikan kesalahan prediksi dalam bentuk persentase yang mudah dipahami dan memungkinkan perbandingan antar model (Makridakis dan Hibon, 1997). RMSE lebih sensitif terhadap kesalahan besar karena memberikan penalti yang lebih tinggi terhadap kesalahan besar dengan menghitung kuadrat perbedaan antara nilai prediksi dan nilai aktual. Hal ini membuat RMSE lebih sesuai untuk mengevaluasi model yang digunakan untuk memprediksi data dengan fluktuasi tinggi, seperti pergerakan IHSG, yang rentan terhadap perubahan drastis (Bates dan Granger, 1969).

MAE mengukur kesalahan keseluruhan tanpa memperhatikan arah kesalahan, sehingga lebih stabil ketika data mengandung outlier atau nilai ekstrim (Theil, 1966). Ketiga metrik ini digunakan untuk menilai dan membandingkan akurasi prediksi antara tiga pendekatan model yaitu Univariate, Multivariate All Feature, dan Multivariate Selected Feature, guna menentukan model yang paling akurat dalam memprediksi pergerakan IHSG. Semakin kecil nilai yang dihasilkan evaluasi metrik maka semakin baik. Selain evaluasi metrik, dilakukan juga uji signifikansi untuk menguji apakah terdapat perbedaan yang signifikan antara kinerja ketiga model yang diuji.

HASIL DAN PEMBAHASAN

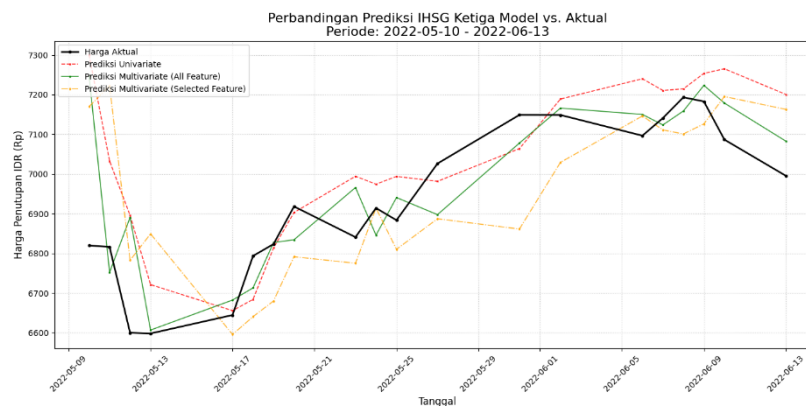
Hasil evaluasi kinerja dari tiga model LSTM, yaitu Univariate, Multivariate All Feature, dan Multivariate Selected Feature. Metrik evaluasi yang digunakan untuk mengukur kinerja ketiga model adalah MAPE, RMSE, dan MAE. Hasil dari masing-masing metrik ini dihitung berdasarkan rata-rata dari 30 kali percobaan menjalankan setiap model. Tabel berikut menyajikan nilai dari ketiga metrik evaluasi untuk masing-masing model:

Tabel 1 Hasil Evaluasi Ketiga Model LSTM

Model LSTM	Evaluasi Metrik		
	MAPE	RMSE	MAE
<i>Univariate</i>	1.28	101.76	86.60
<i>Multivariate All Feature</i>	0.76	66.72	51.58
<i>Multivariate Selected Feature</i>	1.07	94.32	72.35

Dari Tabel 1, terlihat bahwa model LSTM Multivariate All Feature menghasilkan kinerja terbaik di antara ketiga model, dengan nilai MAPE terendah 0.76, diikuti oleh RMSE 66.72, dan MAE 51.58. Hal ini menunjukkan bahwa model ini mampu memprediksi pergerakan IHSG dengan kesalahan yang paling kecil di antara ketiga model. Sementara itu, model LSTM Univariate menunjukkan hasil evaluasi yang paling buruk dengan MAPE tertinggi 1.28, diikuti oleh RMSE 101.76, dan MAE 86.60. Meskipun demikian, model Multivariate Selected Feature juga menunjukkan hasil yang lebih baik dibandingkan dengan Univariate, meskipun tidak sebaik Multivariate All Feature. Hasil ini menunjukkan bahwa penambahan variabel eksternal dalam model Multivariate All Feature memberikan kontribusi signifikan terhadap peningkatan akurasi prediksi.

Untuk visualisasi komparatif yang lebih jelas mengenai perbedaan kinerja ketiga model pada periode pasar yang paling dinamis, dapat dilihat pada Gambar 9. Gambar ini secara spesifik menampilkan perbandingan prediksi masing-masing model dengan pergerakan aktual IHSG pada periode empat minggu yang menunjukkan volatilitas tertinggi dalam data uji, sebagaimana diidentifikasi melalui perhitungan rolling standard deviation dari daily returns.

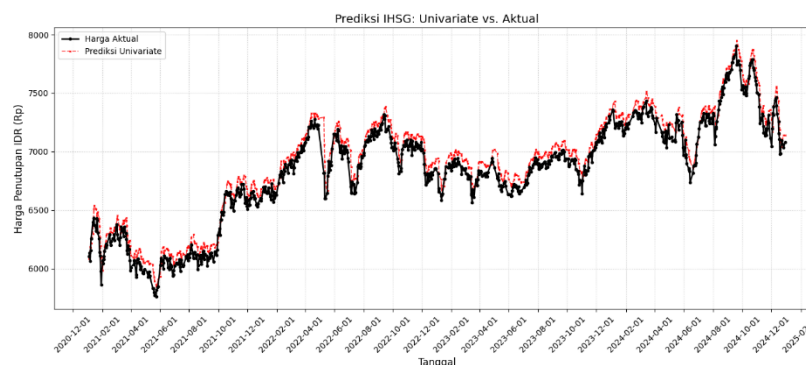


Gambar 1 Perbandingan Prediksi IHSG Berbagai Model LSTM (Periode Paling Fluktuatif)

Untuk memahami lebih dalam bagaimana setiap model memprediksi pergerakan IHSG terhadap pergerakan aktual pada periode training dan testing secara keseluruhan, serta pada periode volatilitas tertinggi, berikut disajikan analisis dan visualisasi untuk masing-masing model.

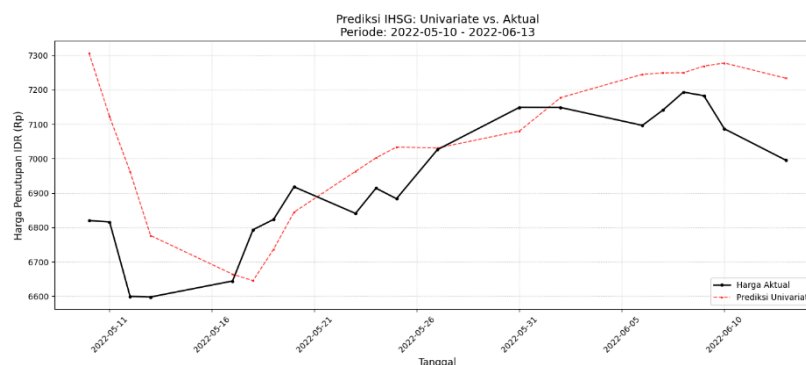
Analisis Model LSTM Univariate

Model LSTM Univariate dilatih hanya menggunakan data historis IHSG sebagai input tunggal. Kinerja umum model ini pada seluruh rentang waktu data (2005-2025) dapat diamati pada Gambar 2.



Gambar 2 Perbandingan LSTM Univariate dengan Pergerakan Aktual

Dari Gambar 2, terlihat bahwa model Univariate secara umum mampu menangkap tren pergerakan IHSG dalam jangka panjang. Namun, model ini cenderung memiliki lag yang signifikan dalam merespons perubahan arah tren yang cepat atau lonjakan harga yang tiba-tiba. Pola prediksi terlihat mengikuti harga aktual, namun dengan penundaan yang cukup jelas, terutama pada periode perubahan harga yang drastis. Untuk detail kinerja yang lebih fokus pada kondisi pasar yang dinamis, perbandingan model Univariate dengan harga aktual pada periode empat minggu paling fluktuatif dapat dilihat pada Gambar 3.

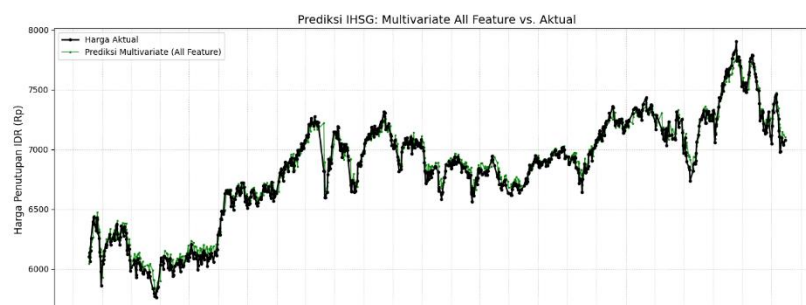


Gambar 3 Perbandingan LSTM Univariate dengan Pergerakan Aktual (Periode Paling Fluktuatif)

Berdasarkan Gambar 3 dan Tabel 1, model Univariate mencatatkan rata-rata MAPE sebesar 1.28%, RMSE 101.76, dan MAE 86.60. Angka-angka ini menunjukkan bahwa meskipun model ini mampu menangkap pola dasar pergerakan IHSG, kinerjanya adalah yang terendah di antara ketiga model yang diuji. Pada periode fluktuatif, garis prediksi model Univariate terlihat sangat tertinggal dan menunjukkan deviasi yang besar dari harga aktual. Hal ini dikarenakan model ini hanya mengandalkan satu variabel (data historis IHSG sebelumnya), sehingga sangat kurang responsif terhadap perubahan dinamis yang dipengaruhi oleh faktor-faktor eksternal pasar saham seperti kondisi ekonomi makro atau sentimen pasar global. Keterbatasan ini menyebabkan model tidak dapat menyesuaikan prediksinya dengan baik di tengah gejolak pasar.

Analisis Model LSTM Multivariate All Feature

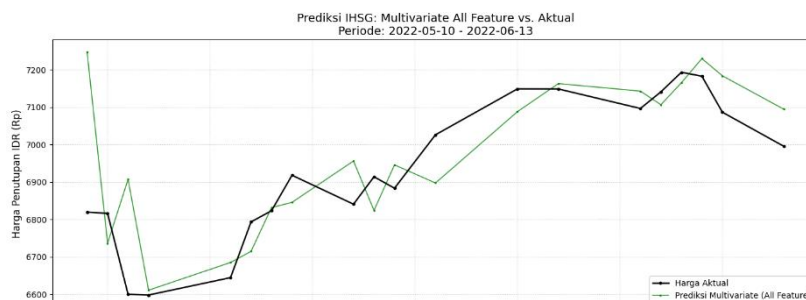
Model LSTM Multivariate All Feature mempertimbangkan berbagai faktor eksternal dalam prediksinya. Kinerja umum model ini pada seluruh rentang waktu data (2005-2025) dapat diamati pada Gambar 4.



Gambar 4 Perbandingan LSTM Multivariate All Feature dengan Pergerakan Aktual

Dari Gambar 4, terlihat bahwa model Multivariate All Feature mampu mengikuti pergerakan IHSG dengan sangat baik di sepanjang periode data yang panjang. Garis prediksi menunjukkan keselarasan yang kuat dengan tren harga aktual, bahkan mampu menangkap beberapa perubahan arah tren yang lebih cepat dibandingkan model Univariate. Hal ini mengindikasikan bahwa penambahan variabel eksternal memberikan model pemahaman yang

lebih komprehensif tentang dinamika pasar. Untuk detail kinerja yang lebih fokus pada kondisi pasar yang dinamis, perbandingan model Multivariate All Feature dengan pergerakan aktual pada periode empat minggu paling fluktuatif dapat dilihat pada Gambar 5.

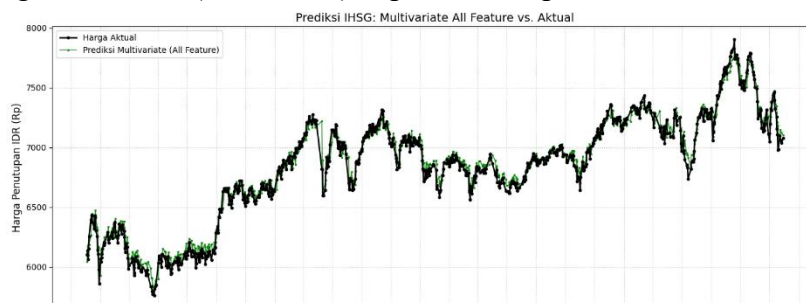


Gambar 5 Perbandingan LSTM Multivariate All Feature dengan Pergerakan Aktual (Periode Paling Fluktuatif)

Berdasarkan Gambar 13 dan Tabel 18, model Multivariate All Feature menunjukkan kinerja terbaik dengan rata-rata MAPE terendah yaitu 0.76%, diikuti oleh RMSE 66.72, dan MAE 51.58. Akurasi yang akurat ini menggarisbawahi kontribusi signifikan dari faktor-faktor eksternal dalam meningkatkan kemampuan prediksi IHSG. Model ini memanfaatkan sejumlah besar variabel relevan seperti Harga Emas Dunia (GC=F), Harga Minyak Dunia (CL=F), Nilai Tukar Rupiah terhadap Dolar Amerika Serikat (IDR=X), Indeks Saham Hong Kong (^HSI), Indeks Saham Jepang (^N225), Indeks Saham Amerika (^DJI), dan Indeks MSCI Emerging Markets (EEM). Dengan menggabungkan berbagai variabel ini, model dapat menangkap pola yang lebih kompleks dalam pergerakan IHSG dan memperhitungkan pengaruh eksternal yang signifikan terhadap pasar saham. Visualisasi pada Gambar 13 secara jelas menunjukkan bahwa garis prediksi model ini sangat mengikuti pergerakan aktual, bahkan pada periode volatilitas tinggi, mengindikasikan kemampuan model untuk membuat prediksi yang lebih dekat dan stabil.

Analisis Model LSTM Multivariate Selected Feature

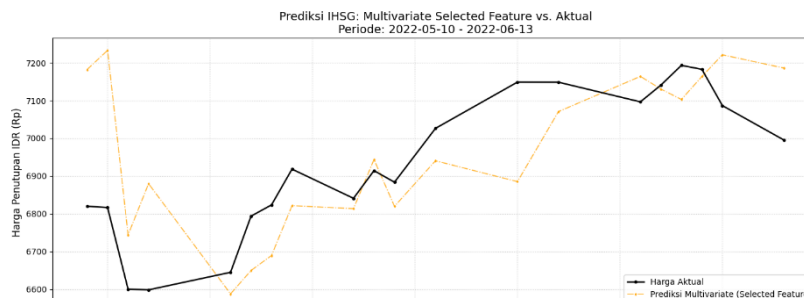
Model LSTM Multivariate Selected Feature dibangun dengan variabel-variabel yang telah diseleksi menggunakan Recursive Feature Elimination (RFE). Kinerja umum model ini pada seluruh rentang waktu data (2005-2025) dapat diamati pada Gambar 6.



Gambar 6 Perbandingan LSTM Multivariate Selected Feature dengan Pergerakan Aktual

Dari Gambar 6, model Multivariate Selected Feature menunjukkan kemampuan yang lebih baik dalam menangkap tren IHSG secara umum dibandingkan model Univariate, yang

menunjukkan bahwa proses feature selection tetap memberikan manfaat dari variabel eksternal. Namun, model ini masih menunjukkan sedikit keterlambatan atau deviasi pada beberapa titik balik pasar dibandingkan dengan model *Multivariate All Feature*. Untuk detail kinerja yang lebih fokus pada kondisi pasar yang dinamis, perbandingan model Multivariate Selected Feature dengan pergerakan aktual pada periode empat minggu paling fluktuatif dapat dilihat pada Gambar 7.



Gambar 7 Perbandingan LSTM Multivariate Selected Feature dengan Pergerakan Aktual (Periode Paling Fluktuatif)

Pada Gambar 7 dan Tabel 1, model Multivariate Selected Feature mencatatkan rata-rata MAPE sebesar 1.07%, RMSE 94.32, dan MAE 72.35. Kinerja model ini terbukti lebih baik dibandingkan model Univariate karena penggunaan variabel eksternal yang relevan. Namun, akurasi masih di bawah model Multivariate All Feature. Hal ini mungkin disebabkan oleh metode feature selection RFE dengan Random Forest yang digunakan. Meskipun feature selection dapat membantu mengurangi dimensi dan menghindari overfitting, beberapa variabel yang terpilih mungkin tidak sepenuhnya cocok dengan karakteristik model LSTM yang sangat bergantung pada urutan waktu dan hubungan kompleks antar variabel dalam data. Oleh karena itu, meskipun variabel yang terpilih relevan, kontribusi kumulatif dari semua variabel dalam model Multivariate All Feature ternyata lebih menguntungkan dan memberikan hasil yang lebih optimal bagi model LSTM dalam memprediksi IHSG.

Analisis Hasil Uji Signifikansi Ketiga Model LSTM

Pada sub-bab ini, dilakukan analisis terhadap hasil uji signifikansi yang telah dilakukan untuk membandingkan kinerja ketiga model LSTM yang diuji, yaitu Univariate, Multivariate All Feature, dan Multivariate Selected Feature. Uji yang digunakan untuk mengevaluasi perbedaan signifikan antar ketiga model adalah Uji ANOVA dan Uji Tukey HSD.

Uji ANOVA menghasilkan nilai F-statistic sebesar 456.176 dan p-value sebesar 0.0, yang jauh lebih kecil dari 0.05. Hal ini menunjukkan bahwa terdapat perbedaan signifikan antara ketiga model yang diuji. Dengan demikian, kita dapat menyimpulkan bahwa kinerja antara model Univariate, Multivariate All Feature, dan Multivariate Selected Feature menunjukkan perbedaan yang signifikan dan tidak bersifat kebetulan, melainkan mencerminkan perbedaan nyata dalam kemampuan masing-masing model dalam melakukan prediksi.

Setelah uji ANOVA menunjukkan adanya perbedaan signifikan, dilakukan uji Tukey HSD untuk mengetahui pasangan model mana yang menunjukkan perbedaan yang signifikan. Berikut adalah hasil uji Tukey HSD untuk pasangan-pasangan model:

Multivariate All Feature dengan Multivariate Selected Feature: Nilai p-adj adalah 0.0, yang lebih kecil dari 0.05, menunjukkan bahwa perbedaan antara kedua model ini signifikan.

Multivariate All Feature dengan Univariate: Nilai p-adj adalah 0.0, yang juga lebih kecil dari 0.05, menunjukkan bahwa perbedaan antara kedua model ini signifikan.

Multivariate Selected Feature dengan Univariate: Nilai p-adj adalah 0.0, yang lebih kecil dari 0.05, menunjukkan bahwa perbedaan antara kedua model ini juga signifikan.

Hasil uji ANOVA menunjukkan adanya perbedaan signifikan antara ketiga model, yang kemudian diperjelas dengan uji Tukey HSD. Setiap pasangan model yang dibandingkan menunjukkan perbedaan yang signifikan secara statistik, dengan nilai p-adj yang sangat kecil. Ini menunjukkan bahwa Multivariate All Feature memiliki kinerja yang lebih baik dibandingkan dengan model Univariate dan Multivariate Selected Feature, dan kedua model Multivariate juga memiliki perbedaan yang signifikan satu sama lain. Secara keseluruhan, hasil uji signifikansi ini mengkonfirmasi bahwa Multivariate All Feature memberikan kinerja yang paling baik dalam memprediksi pergerakan IHSG dibandingkan dengan model lainnya.

KESIMPULAN

Model Prediksi Univariate Menggunakan LSTM: Model ini menggunakan data historis IHSG sebagai input tunggal, dengan hasil evaluasi MAPE 1.28, RMSE 101.76, dan MAE 86.60. Meskipun dapat memprediksi pergerakan IHSG, hasilnya kurang optimal dibandingkan model Multivariate. Model Prediksi Multivariate All Feature Menggunakan LSTM: Model ini mempertimbangkan berbagai faktor eksternal seperti harga emas, minyak, dan nilai tukar. Hasil evaluasi menunjukkan MAPE 0.76, RMSE 66.72, dan MAE 51.58, yang menjadikannya model terbaik di antara ketiganya. Model Prediksi Multivariate Selected Feature Menggunakan LSTM dengan Feature Selection: Menggunakan feature selection dengan Random Forest, model ini menunjukkan MAPE 1.07, RMSE 94.32, dan MAE 72.35, namun kinerjanya masih di bawah Multivariate All Feature. Perbandingan Hasil Error dari Ketiga Model: Model Multivariate All Feature memiliki error terkecil dengan MAPE 0.76, RMSE 66.72, dan MAE 51.58, sementara Multivariate Selected Feature lebih baik daripada Univariate, tetapi tidak seakurat Multivariate All Feature.

DAFTAR PUSTAKA

- Bareksa. (2024). *Rebalancing MSCI dorong asing nett sell Rp1,46 T, IHSG turun 2,5%*. Diakses dari <https://www.bareksa.com/berita/pasar-modal/2015-11-01/rebalancing-msci-dorong-asing-nett-sell-rp146-t-ihsg-turun-25>
- Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4), 451–468. <https://doi.org/10.1057/jors.1969.92>
- Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.
- Chabachib, H. M., & Witjaksono, A. A. (2011). *Analisis pengaruh fundamental makro dan indeks harga global terhadap IHSG*. Universitas Indonesia.
- Gao, T., Chai, Y., & Liu, Y. (2018). Applying long short-term memory neural networks for predicting stock closing price. *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences (ICSESS)*, 575–578. <https://doi.org/10.1109/ICSESS.2017.8342981>

- Gumilang, R. C. (2014). *Pengaruh variabel makro ekonomi, harga emas dan harga minyak dunia terhadap Indeks Harga Saham Gabungan (Studi pada Bursa Efek Indonesia periode 2009–2013)*. Skripsi, Fakultas Ilmu Administrasi, Universitas Brawijaya.
- Hassan, M. K. (2019). Predicting stock market movements: A study of the Indonesian stock market. *International Journal of Economics and Financial Issues*, 9(1), 85–92. <https://doi.org/10.2139/ssrn.3342385>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Idrees, S. M., Alam, M. A., & Agarwal, P. (2019). A prediction approach for stock market volatility based on time series data. *IEEE Access*, 7, 17287–17298. <https://doi.org/10.1109/ACCESS.2019.2895252>
- Invesnow. (2024). *MSCI reduces Indonesia's weight in Emerging Markets Index: Impact on IHSG*. Diakses dari <https://invesnow.id/artikel/msci-reduces-indonesias-weight>
- Krauss, C., Do, X. A., & Huck, N. (2017). *Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500*. FAU Discussion Papers in Economics, 03/2016.
- Madi, I. (2018). Forecasting stock market index using ARIMA and LSTM. *Journal of Financial Studies*, 12(4), 55–68. <https://doi.org/10.1080/21427844.2018.1456781>
- Makridakis, S., & Hibon, M. (1997). The accuracy of forecasting: An evaluation of 120 forecasting methods. *International Journal of Forecasting*, 5(4), 455–469. [https://doi.org/10.1016/0169-2070\(89\)90022-3](https://doi.org/10.1016/0169-2070(89)90022-3)
- Putra, D. A., & Nurmatias. (2024). Analisis Indeks Harga Saham Global dan Inflasi Terhadap Indeks Harga Saham Gabungan (IHSG). *IKRAITH-EKONOMIKA*, 7(1), 93–101. <https://doi.org/10.37817/ikraith-ekonomika.v7i1>
- Sezer, M., Gudelek, M. I., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review (2005–2019). *Applied Soft Computing*, 90, 106181. <https://doi.org/10.1016/j.asoc.2020.106181>
- Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2019). A comparison of ARIMA and LSTM in forecasting time series. *Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA 2018)*, 1394–1401. <https://doi.org/10.1109/ICMLA.2018.00227>
- Theil, H. (1966). *Applied economic forecasting*. Rand McNally.