



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 4 (2025) pp: 829-837

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Perbandingan Kinerja Model Machine Learning dalam Prediksi Harga Saham TLKM

Suyatna, Adi Joko Satrio, Satrio Putra Pratama, Mamat Indarto, Fadillah Syarif, Fachry Amsury

Teknologi Informasi, Teknik dan Informatika, Universitas Bina Sarana Informatika

suyatna644@gmail.com, adijokos21@gmail.com, satrioputra204@gmail.com, soka.ubsi@gmail.com,

fadilisyarif378@gmail.com, fachri.fey@bsi.ac.id

Abstrak

Penelitian ini bertujuan mengevaluasi tingkat kinerja model machine learning dalam memprediksi harga penutupan saham TLKM berdasarkan data historis periode 2005 hingga 2024. Tiga algoritma diterapkan, yaitu linear regression, random forest, dan neural network (MLP), yang masing-masing mewakili tingkat kompleksitas berbeda. Evaluasi dilakukan menggunakan teknik cross validation dengan tiga metrik utama, yakni R^2 , MAE, dan RMSE, untuk menilai ketepatan serta stabilitas hasil prediksi. Hasil penelitian menunjukkan bahwa random forest memiliki performa paling unggul dibandingkan dua model lainnya karena mampu menghasilkan prediksi yang stabil, akurat, dan efisien pada berbagai variasi data. Model ini juga menunjukkan kemampuan generalisasi yang tinggi, di mana hasil prediksi pada data uji memiliki tingkat kesesuaian hampir sempurna dengan nilai aktual. Linear regression memberikan hasil yang baik dan efektif dalam mengenali pola hubungan linier antarvariabel, mencerminkan bahwa data saham TLKM memiliki kecenderungan hubungan yang relatif stabil antar atributnya. Neural network (MLP) memperlihatkan peningkatan kinerja setelah dilakukan hyperparameter tuning, terutama pada konfigurasi jumlah neuron dan lapisan tersembunyi yang tepat, namun efisiensi komputasi model ini masih lebih rendah karena proses pelatihannya memerlukan waktu yang lebih panjang serta sensitif terhadap perubahan parameter. Keseluruhan, Random forest dinilai sebagai model paling optimal karena mampu menjaga keseimbangan antara akurasi, stabilitas, dan efisiensi komputasi. Temuan ini memperkuat potensi penerapan algoritma ensemble dalam analisis pasar modal sebagai pendekatan yang lebih adaptif terhadap data keuangan. Hasil penelitian diharapkan menjadi referensi bagi mahasiswa, peneliti, dan investor dalam mengembangkan model prediksi berbasis machine learning yang andal untuk mendukung pengambilan keputusan investasi berbasis data di pasar saham Indonesia.

Kata kunci: Prediksi Harga Saham, Machine Learning, Linear Regression, Random Forest, Neural Network, Cross Validation

1. Pendahuluan

Saham adalah instrumen investasi yang berperan tidak hanya sebagai media penanaman modal, tetapi juga sebagai sumber pembiayaan yang efisien bagi perusahaan, sehingga peran saham sangat strategis dalam mendorong pertumbuhan perekonomian suatu negara [1]. Pada beberapa tahun terakhir, minat masyarakat Indonesia terhadap investasi saham menunjukkan peningkatan yang cukup signifikan [2]. Berdasarkan data pasar modal, jumlah investor saham meningkat dari 3.451.513 orang pada tahun 2021 menjadi 5.255.571 orang pada tahun 2023 [3]. Peningkatan disebabkan oleh pandangan bahwa investasi saham mampu memberikan keuntungan lebih tinggi dibandingkan instrumen investasi lainnya [4]. Sementara, pergerakan harga saham yang tergolong fluktuatif menghadirkan risiko cukup besar bagi investor [5]. Atas dasar tersebut, penelitian terkait prediksi harga saham tetap menjadi topik yang relevan dan banyak diteliti hingga saat ini [6].

PT Telkom Indonesia (Persero) Tbk merupakan salah satu Badan Usaha Milik Negara (BUMN) yang beroperasi di bidang telekomunikasi dan memiliki peranan penting dalam penyediaan layanan komunikasi di Indonesia [7]. Saham dengan kode emiten TLKM tergolong dalam kategori *blue chip* serta termasuk di antara saham yang paling aktif diperdagangkan di Bursa Efek Indonesia (BEI) [8]. Tingginya aktivitas perdagangan menunjukkan besarnya minat investor terhadap saham ini. Pergerakan harga penutupan saham TLKM masih memperlihatkan tingkat fluktuasi yang cukup tinggi [9]. Kondisi tersebut dipengaruhi oleh sejumlah faktor, baik dari aspek internal perusahaan seperti kondisi laporan keuangan, maupun dari faktor eksternal seperti kondisi ekonomi makro dan dinamika pasar nasional [10].

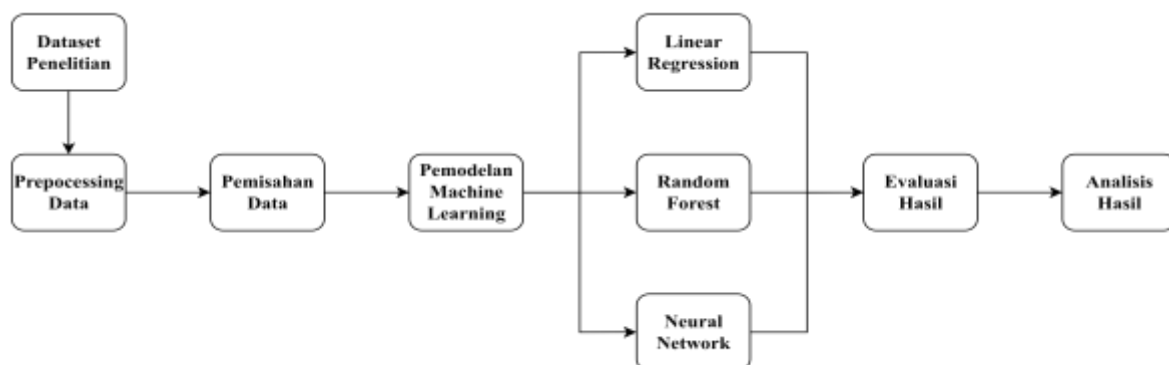
Analisis terhadap pergerakan saham umumnya dilakukan melalui dua pendekatan utama, yaitu analisis fundamental dan analisis teknikal [11]. Pendekatan fundamental berfokus pada aspek internal perusahaan, seperti kondisi keuangan, tingkat profitabilitas, serta prospek bisnis di masa mendatang. Berbeda dengan pendekatan fundamental, analisis teknikal memusatkan analisis pada identifikasi pola pergerakan harga saham dengan memanfaatkan data historis sebagai dasar pengambilan keputusan investasi [12]. Analisis fundamental dan analisis teknikal telah lama digunakan secara luas oleh investor dan analis pasar modal dalam menilai potensi kinerja saham [13]. Seiring dengan perkembangan teknologi di bidang *data science*, muncul pendekatan baru melalui penerapan *machine learning* yang memberikan cara pandang berbeda dalam analisis data keuangan [14]. Pendekatan ini memungkinkan sistem komputer untuk mempelajari pola dari data yang tersedia serta mengenali hubungan nonlinier yang sulit diungkap melalui metode analisis konvensional [15].

Model *machine learning* terbukti mampu memberikan hasil prediksi yang cukup akurat dalam menganalisis data keuangan [16]. Tingkat kinerja setiap model tidak selalu seragam karena sangat bergantung pada karakteristik dataset yang digunakan [17]. Sejumlah penelitian perbandingan menunjukkan bahwa model seperti *convolutional neural network* dan *gated recurrent unit* memiliki keunggulan dalam mengidentifikasi pola harga saham dengan tingkat fluktuasi yang tinggi [18]. Berbeda dengan model *long short term memory* dan *recurrent neural network* yang cenderung tidak optimal ketika diterapkan pada data dengan pergerakan harga yang tidak stabil [19]. Berdasarkan hal tersebut, pengujian terhadap beberapa model perlu dilakukan pada kasus saham TLKM guna memperoleh metode yang paling tepat serta mampu menghasilkan prediksi harga yang akurat dan konsisten.

Penelitian ini memanfaatkan data historis harga saham TLKM dalam rentang waktu 2005 hingga 2024. Data diolah menggunakan tiga algoritma *machine learning*, yaitu *linear regression*, *random forest*, dan *neural network* (MLP). Ketiga model tersebut dipilih untuk merepresentasikan tingkat kompleksitas yang berbeda, mulai dari model sederhana, menengah, hingga kompleks [20]. Seluruh proses *preprocessing*, hingga evaluasi dan analisis hasil dilakukan dengan menggunakan perangkat lunak Orange Data Mining, yang mendukung pendekatan analisis berbasis visual *workflow* sehingga memudahkan penerapan berbagai algoritma *machine learning* secara terstruktur dan efisien. Melalui tahap evaluasi dan perbandingan kinerja, penelitian ini bertujuan untuk mengidentifikasi model yang paling akurat dalam memprediksi harga penutupan saham TLKM [21]. Hasil penelitian diharapkan dapat menjadi referensi bagi mahasiswa, peneliti, maupun investor yang tertarik pada penerapan *machine learning* dalam bidang analisis keuangan.

2. Metode Penelitian

Penelitian yang dilakukan, menggunakan pendekatan kuantitatif dengan memanfaatkan data historis harga saham TLKM sebagai objek analisis. Seluruh tahapan penelitian yang meliputi *processing*, pemisahan data, pemodelan *machine learning*, serta evaluasi dan analisis hasil, dilaksanakan menggunakan perangkat lunak Orange Data Mining. Pemilihan Orange Data Mining didasari pada kemampuannya, yang menyediakan antarmuka berbasis *visual programming* yang memungkinkan pengguna membangun alur kerja analisis *machine learning* secara interaktif, efisien, dan terstruktur. Melalui pendekatan berbasis alur kerja tersebut, setiap tahapan mulai dari *preprocessing*, pemisahan data, pemodelan, hingga evaluasi dan analisis dapat dijalankan secara sistematis dan terkontrol, sehingga mendukung validitas hasil analisis yang dihasilkan.



Gambar 1. Tahapan Metode Penelitian

2.1. Dataset Penelitian

Dataset yang digunakan dalam penelitian ini merupakan data historis harga saham TLKM periode 2005 hingga 2024, yang terdiri dari 4705 baris. Rentang waktu dipilih untuk memberikan cakupan observasi yang lebih luas sehingga pola fluktuasi harga saham dapat dianalisis secara komprehensif. Dataset diperoleh dari sumber publik Kaggle, yang telah melalui proses pembersihan dan standarisasi sehingga siap digunakan untuk analisis lanjutan. Dataset diunduh dalam format .csv dan mencakup beberapa atribut, yaitu *Open*, *High*, *Low*, *Close*, *Adj Close*, *Volume*, dan *Date*. Fokus utama penelitian ini adalah memprediksi harga penutupan (*Close*). Atribut *Open*, *High*, *Low*, dan *Volume* digunakan sebagai variabel input, kecuali variabel *Date* yang hanya berfungsi sebagai penanda waktu dan tidak dilibatkan secara langsung, sehingga potensi kebocoran data dapat dihindari.

Tabel 1. Data TLKM Periode 2005 - 2024

Open	High	Low	Close	Adj Cose	Volume	Date
965.000000	975.000000	965.000000	970.000000	548.346130	84725000	2005-01-26
970.000000	970.000000	960.000000	960.000000	542.692932	134222500	2005-01-27
965.000000	965.000000	955.000000	960.000000	542.692932	100760000	2005-01-28
960.000000	965.000000	950.000000	960.000000	542.692932	137322500	2005-01-31
...

2.2. Preprocessing Data

Tahap *preprocessing* merupakan proses awal yang memiliki peran penting dalam memastikan kualitas data sebelum digunakan dalam pemodelan. Pada tahap *preprocessing* dilakukan serangkaian langkah pembersihan dan penyesuaian data untuk meningkatkan akurasi hasil prediksi. Proses dimulai dengan pemeriksaan terhadap adanya *missing values* serta identifikasi format data yang tidak sesuai. Hasil observasi menunjukkan bahwa beberapa variabel mengandung nilai kosong, sehingga dilakukan penanganan menggunakan widget *impute* dengan metode pengisian berdasarkan nilai rata-rata. Selain itu, ditemukan bahwa variabel *Open*, *High*, *Low*, *Close*, *Volume*, dan *Adj Close* terbaca sebagai tipe *text*, bukan *numeric*. Permasalahan tersebut diperbaiki dengan melakukan konversi tipe data menjadi *numeric*, supaya dapat diolah secara komputasional. Usai proses konversi dilakukan, variabel *Close* ditetapkan sebagai target variabel yang akan digunakan dalam proses prediksi.

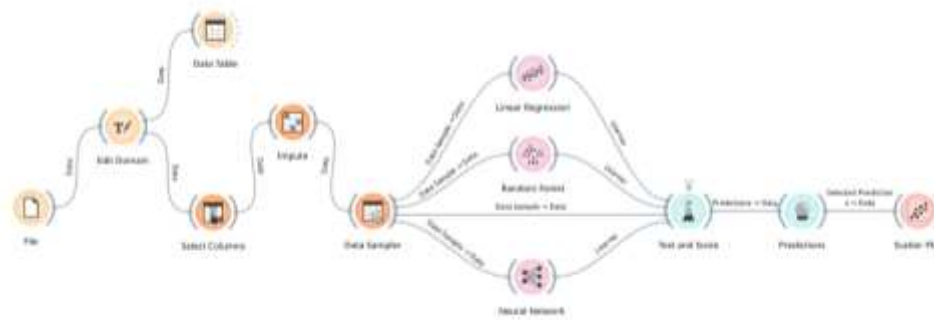
Tabel 2. Data Saham Setelah Preprocessing Data

Open	High	Low	Close	Adj Cose	Volume	Date
965	975	965	970	548.346	8,4725e+07	2005-01-26
970	970	960	960	542.693	1,34222e+08	2005-01-27
965	965	955	960	542.693	1,0076e+08	2005-01-28
960	965	950	960	542.693	1,37322e+08	2005-01-31
...

2.3. Pemisahan Data

Selesai proses pembersihan data, tahap berikutnya adalah pembagian dataset untuk keperluan pelatihan dan pengujian model. Proses ini dilakukan menggunakan widget *data sampler* dengan proporsi pembagian sebesar 80% untuk data latih dan sisanya digunakan untuk data *testing*. Pembagian tersebut bertujuan supaya proses pelatihan model dapat berjalan secara efisien sekaligus memberikan ruang evaluasi terhadap kemampuan model dalam menghadapi data baru. Selain itu, digunakan metode *cross validation* dengan jumlah lipatan sebanyak sepuluh, di mana setiap lipatan secara bergantian berperan sebagai data *testing*, sementara sembilan lipatan lainnya digunakan untuk pelatihan. Metode *cross validation* diterapkan berguna untuk memastikan bahwa model tidak hanya memiliki performa yang baik pada data latih, tetapi juga mampu melakukan generalisasi secara konsisten terhadap data yang belum pernah dilihat sebelumnya.

2.4. Pemodelan Machine Learning



Gambar 2. Alur Kerja Machine Learning

Pada tahap pemodelan *machine learning*, model *linear regression*, *random forest*, serta *neural network* (MLP) awalnya diterapkan dengan konfigurasi *default*, kemudian dikembangkan melalui *hyperparameter tuning* dengan tujuan meningkatkan kinerja prediksi. *Linear regression* dikembangkan melalui penerapan berbagai variasi, berupa *lasso*, *ridge*, dan *elastic net*, dengan pengujian beberapa nilai parameter α (*alpha*) serta *ratio_l1* untuk mengevaluasi pengaruh regularisasi terhadap performa model. Selain *linear regression*, *random forest* juga dikembangkan dengan variasi jumlah pohon (*n_estimators*) dan kedalaman maksimum (*max_depth*) untuk memperoleh keseimbangan antara akurasi dan risiko *overfitting*, sedangkan *neural network* (MLP) diuji dengan penyesuaian jumlah lapisan tersembunyi (*hidden_layer*) dan batas iterasi maksimum (*max_iter*) untuk meningkatkan kemampuan mengenali pola nonlinier. Melalui kombinasi berbagai algoritma dan proses *hyperparameter tuning*, penelitian ini bertujuan memperoleh model yang paling optimal dalam memprediksi harga penutupan saham TLKM.

Tabel 3. Model Machine Learning dengan Konfigurasi Default

Model	Hyperparameter (A)	Nilai Diuji (A)	Hyperparameter (B)	Nilai Diuji (B)
Linear regression	α (alpha)	0.0001	l1_ratio	0.50
Random forest	n_estimators	100	max_depth	3
Neural network (MLP)	hidden_layer	[10, 10]	max_iter	300

Tabel 4. Variasi Hyperparameter Tuning Linear Regression

Model	Hyperparameter	Nilai Diuji
Lasso regression (L1)	α (alpha)	0.1, 1, 10
Ridge regression (L2)	α (alpha)	0.1, 1, 10
Elastic net regression (L1 + L2)	α (alpha), l1_ratio	$\alpha = 0.1, 1, 10$ ratio_l1 = 0.30, 0.50, 0.70

Tabel 5. Variasi Hyperparameter Tuning Random Forest

Model	Hyperparameter	Nilai Diuji
Random forest	n_estimators	50, 100, 200
	max_depth	10, 20

Tabel 6. Variasi Hyperparameter Tuning Neural Network (MLP)

Model	Hyperparameter	Nilai Diuji
Neural network (MLP)	hidden_layer	[50], [100], [50,50], [100,100]
	max_iter	300

2.5. Evaluasi dan Analisis Hasil

Model yang telah selesai dilatih, selanjutnya dievaluasi menggunakan sejumlah metrik regresi untuk mengukur tingkat akurasi hasil prediksi. Penelitian ini memanfaatkan tiga indikator utama, yaitu R^2 , MAE, dan RMSE. Nilai R^2 digunakan untuk menilai seberapa besar proporsi variasi data yang dapat dijelaskan oleh model, sedangkan MAE dan RMSE berfungsi mengukur rata-rata selisih antara nilai prediksi dengan nilai aktual. Model yang memperoleh nilai R^2 tertinggi serta memiliki nilai MAE dan RMSE terendah dianggap sebagai model dengan performa terbaik. Selain evaluasi berdasarkan parameter bawaan, dilakukan pula pengujian lanjutan setelah proses *hyperparameter tuning* guna mengetahui pengaruh optimasi terhadap peningkatan kinerja model. Sebagai penunjang analisis, hasil evaluasi juga divisualisasikan melalui *scatter plot* untuk memperlihatkan tingkat kesesuaian antara nilai aktual dan hasil prediksi yang dihasilkan oleh masing-masing model.

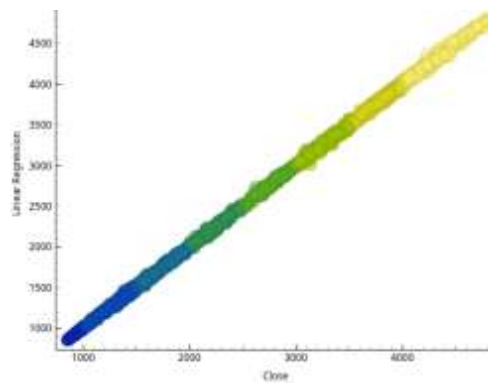
3. Hasil dan Diskusi

3.1. Hasil Evaluasi

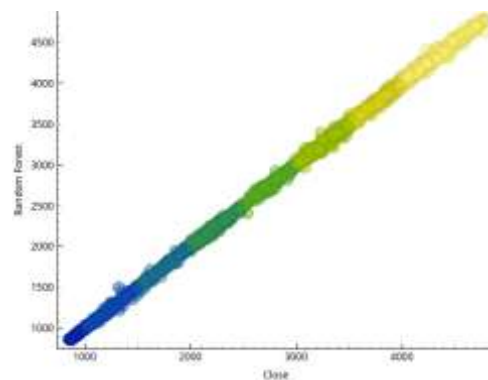
Pada penelitian ini dilakukan perbandingan terhadap tiga algoritma *machine learning* untuk memprediksi harga penutupan saham TLKM. Seluruh model diuji menggunakan metode *10 fold cross validation* melalui widget *test and score* pada perangkat lunak Orange Data Mining. Evaluasi kinerja model dilakukan dengan mengacu pada tiga metrik utama, yaitu R^2 , MAE, dan RMSE. Setiap algoritma dilakukan dengan proses *hyperparameter tuning* untuk memperoleh kombinasi parameter yang paling optimal. Variasi pengujian meliputi penerapan berupa nilai regulasi pada *linear regression*, penyesuaian jumlah pohon dan kedalaman maksimum pada *random forest*, serta pengaturan jumlah layer tersembunyi dan batas iterasi maksimum pada *neural network* (MLP). Tabel berikut adalah hasil dan evaluasi berupa visual *scatter plot* model tanpa dilakukan proses *hyperparameter tuning*.

Tabel 7. Hasil Evaluasi Model Tanpa Hyperparameter Tuning

Model	R^2	MAE	RMSE
Linear regression	1.000	16.706	22.478
Random forest	1.000	15.511	24.241
Neural network (MLP)	0.948	180.255	252.722



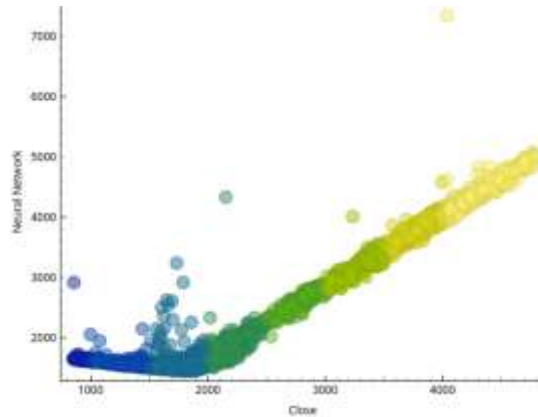
Gambar 3. Visual Scatter Plot Linear Regression



Gambar 4. Visual Scatter Plot Random Forest

DOI: <https://doi.org/10.31004/riggs.v4i4.3476>

Lisensi: Creative Commons Attribution 4.0 International (CC BY 4.0)



Gambar 5. Visual Scatter Plot Neural Network (MLP)

Berdasarkan hasil pengujian awal, model *linear regression* dan *random forest* menunjukkan nilai R^2 sebesar 1.000. Nilai ini menandakan bahwa hampir seluruh variasi pada data harga penutupan saham dapat dijelaskan secara sempurna oleh kedua model tersebut. Berbeda dengan sebelumnya, model *neural network* (MLP) memperoleh nilai R^2 sebesar 0.948, hasil tersebut masih mencerminkan kemampuan prediksi yang cukup tinggi, meskipun sedikit di bawah dua model lainnya. Pada tingkat kesalahan, *random forest* menghasilkan nilai MAE paling rendah, yaitu sebesar 15.511, yang mengindikasikan bahwa selisih rata-rata antara nilai aktual dan hasil prediksi relatif kecil. Dari sudut lain, *linear regression* mencatat nilai RMSE terendah sebesar 22.478, menunjukkan bahwa model ini memiliki kestabilan prediksi yang baik tanpa ada penyimpangan besar. Terdapat juga *neural network* (MLP) yang memperlihatkan nilai MAE dan RMSE yang lebih tinggi dibandingkan model lainnya, sehingga dapat artikan bahwa algoritma ini kurang optimal dalam mengenali pola data saham TLKM yang cenderung bersifat linier.

3.2. Analisis Performa

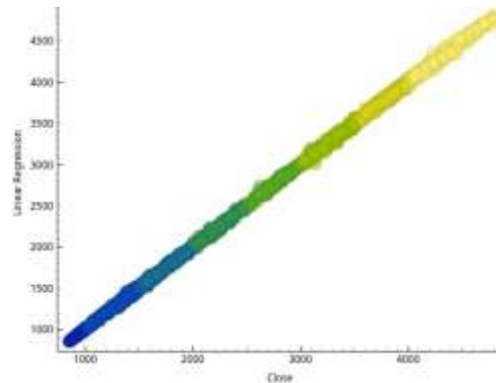
Model *linear regression* diterapkan sebagai *baseline* model untuk menganalisis hubungan linear antara variabel *Open*, *High*, *Low*, dan *Volume* terhadap variabel *Close*. Hasil evaluasi awal menunjukkan bahwa model ini memiliki kinerja yang sangat baik, dengan nilai R^2 sebesar 1.000, MAE sebesar 16.706, dan RMSE sebesar 22.478. Nilai tersebut mengindikasikan bahwa rata-rata kesalahan prediksi model hanya berada pada kisaran 16 hingga 22 poin dari nilai aktual. Model *linear regression* yang sederhana telah mampu menangkap pola hubungan antarvariabel secara akurat tanpa memerlukan struktur model yang kompleks. Tahap selanjutnya, dilakukan proses *hyperparameter tuning* dengan menambahkan variasi regularisasi, meliputi *lasso regression*, *ridge regression*, dan *elastic net regression*, serta pengujian terhadap beberapa nilai parameter α (*alpha*) dan *ratio_l1* untuk menilai pengaruh regularisasi terhadap performa model.

Tabel 8. Analisis Performa Model Linier Regression

Model	α (alpha)	l1_ratio	R^2	MAE	RMSE
Lasso regression (L1)	0.1	-	0.999	19.958	30.407
	1	-	0.999	19.957	30.404
	10	-	0.999	19.952	30.395
Ridge regression (L2)	0.1	-	1.000	16.706	22.478
	1	-	1.000	16.706	22.478
	10	-	1.000	16.706	22.478
Elastic net regression (L1 + L2)	0.1	0.30	0.999	19.958	30.407
	1	0.50	0.999	19.956	30.402
	10	0.70	0.999	19.944	30.375

Analisis terhadap hasil pengujian mengindikasikan, bahwa penerapan *lasso regression* dengan beberapa variasi nilai α (*alpha*) menghasilkan nilai R^2 sebesar 0.999, MAE berkisar antara 19.958 hingga 19.952, serta RMSE antara 30.407 hingga 30.395. Nilai tersebut sedikit lebih rendah dibandingkan model dasar *linear regression*, sementara *ridge regression* dengan beberapa variasi α (*alpha*) menghasilkan performa yang serupa, dengan nilai R^2 sebesar 1.000, MAE sebesar 16.706, dan RMSE sebesar 22.478. Adapun *elastic net regression* yang memberikan hasil relatif sama dengan *lasso regression*, yaitu dengan nilai R^2 sebesar 0.999, MAE berkisar antara 19.958 hingga 19.944, serta RMSE antara 30.407 hingga 30.375. Temuan ini menunjukkan bahwa hubungan

antarvariabel pada dataset bersifat sangat linier, sehingga penerapan regularisasi tidak memberikan pengaruh signifikan terhadap peningkatan performa model



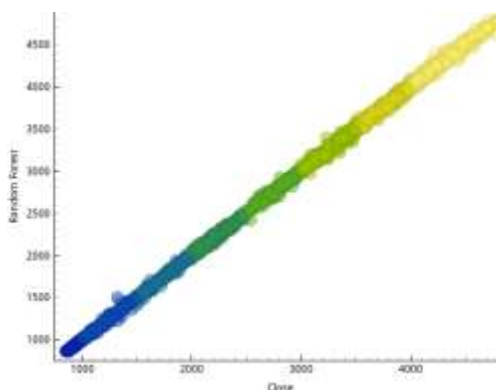
Gambar 6. Visual Scatter Plot Model Default Linear Regression

Berbeda dengan model sebelumnya, model *random forest* diterapkan sebagai algoritma *ensemble* yang menggabungkan sejumlah pohon keputusan guna meningkatkan stabilitas serta akurasi hasil prediksi. Hasil pengujian awal, menunjukkan bahwa model *random forest* memiliki kinerja yang sangat baik, dengan menghasilkan nilai R^2 sebesar 1.000, MAE sebesar 15.511, dan RMSE sebesar 24.241. Nilai yang dihasilkan, mengindikasikan bahwa model *random forest* memiliki performa sedikit lebih unggul dibandingkan model bawaan *linear regression*, meskipun perbedaannya relatif kecil sekitar 15 hingga 24 poin. Proses *hyperparameter tuning* dilakukan untuk mengamati pengaruh terhadap peningkatan performa model dengan mengatur sejumlah pohon ($n_estimators$) dan kedalaman maksimum (max_depth).

Tabel 9. Analisis Performa Model Random Forest

Model	$n_estimators$	max_depth	R^2	MAE	RMSE
Random forest	50	10	1.000	15.504	24.218
	100	10	1.000	15.451	24.125
	200	10	1.000	15.445	24.069
	50	20	1.000	15.602	24.376
	100	20	1.000	15.508	24.272
	200	20	1.000	15.463	24.206

Berdasarkan hasil pengujian yang diperoleh, model *random forest* dengan variasi jumlah pohon dan kedalaman maksimum menunjukkan performa yang konsisten dan stabil pada seluruh kombinasi parameter. Nilai R^2 tercatat sebesar 1.000 di setiap konfigurasi, yang menandakan bahwa model mampu menjelaskan hampir seluruh variasi pada data harga penutupan saham TLKM secara akurat. Nilai MAE berada pada kisaran 15.602 hingga 15.445, sedangkan RMSE berkisar antara 24.376 hingga 24.069. Kombinasi parameter terbaik diperoleh pada jumlah pohon ($n_estimators$) sebesar 200 dan maksimum kedalaman (max_depth) sebesar 10, dengan MAE sebesar 15.445 serta RMSE sebesar 24.069. Hasil tersebut menunjukkan bahwa peningkatan jumlah pohon dalam model mampu memperkuat stabilitas prediksi tanpa menimbulkan kenaikan tingkat kesalahan yang signifikan.



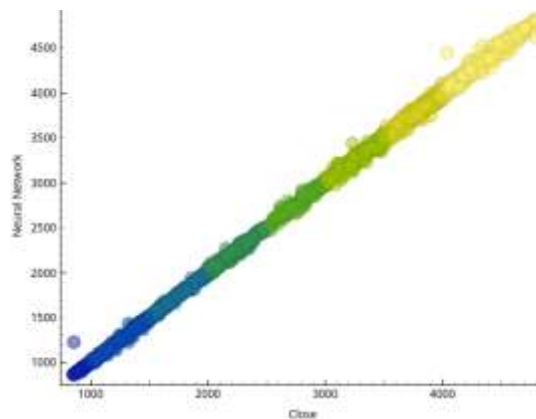
Gambar 7. Visual Scatter Plot Random Forest dengan Kombinasi Hyperparameter Terbaik

Pada penelitian ini, model *neural network* (MLP) juga digunakan sebagai algoritma nonlinier yang memiliki kemampuan untuk mengenali pola kompleks melalui proses propagasi antar neuron. Hasil pengujian awal sebelum dilakukan *hyperparameter tuning* menunjukkan bahwa performa model masih terbatas, dengan nilai R^2 sebesar 0.948, MAE sebesar 180.255, dan RMSE sebesar 252.722. Nilai tersebut menunjukkan bahwa model *neural network* (MLP) belum mampu sepenuhnya menangkap hubungan antarvariabel secara akurat, terutama pada data saham TLKM yang memiliki kecenderungan pola linier. Proses *hyperparameter tuning* tersebut dilakukan untuk menguji pengaruh variasi jumlah lapisan tersembunyi (*hidden_layer*) dan jumlah iterasi maksimum (*max_iter*) terhadap peningkatan performa model.

Tabel 10. Analisis Performa Model Neural Network (MLP)

Model	<i>hidden_layer</i>	<i>max_iter</i>	R^2	MAE	RMSE
Neural network (MLP)	[50]	300	0.526	614.725	765.663
	[100]	300	0.729	465.522	579.203
	[50, 50]	300	0.999	20.312	32.928
	[100, 100]	300	0.999	19.051	28.914

Temuan dari pengujian model menunjukkan, bahwa semakin banyak lapisan tersembunyi (*hidden_layer*) yang digunakan, maka semakin baik juga kemampuan model dalam mengenali hubungan nonlinier antarvariabel input. Secara keseluruhan, model *neural network* (MLP) dengan konfigurasi dua lapisan tersembunyi (*hidden_layer*) sebanyak [100, 100] menghasilkan performa terbaik dibandingkan variasi lainnya, dengan tingkat akurasi yang sebanding dengan beberapa variasi model *random forest* dan *linear regression* setelah melalui proses *hyperparameter tuning*. Dibandingkan dengan dua model tersebut, *neural network* (MLP) memerlukan waktu pelatihan yang lebih lama serta menunjukkan sensitivitas yang lebih tinggi terhadap perubahan parameter. Artinya, *neural network* (MLP) unggul dalam mengenali pola data yang kompleks, penerapannya tetap perlu mempertimbangkan keseimbangan antara tingkat akurasi dan efisiensi komputasi.



Gambar 8. Visual Scatter Plot Neural Network dengan Kombinasi Hyperparameter Terbaik

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan terhadap tiga algoritma *machine learning*, yaitu *linear regression*, *random forest*, dan *neural network* (MLP), dapat disimpulkan bahwa penerapan model *machine learning* mampu memberikan hasil prediksi yang sangat akurat terhadap harga penutupan saham TLKM. Berdasarkan hasil evaluasi, *random forest* terbukti sebagai model dengan performa paling unggul dan stabil, di mana pada konfigurasi terbaik dengan jumlah pohon (*n_estimators*) sebanyak 200 dan maksimum kedalaman (*max_depth*) sebesar 10 yang menunjukkan bahwa model mampu menjelaskan hampir seluruh variasi data dengan tingkat kesalahan yang sangat kecil. Model *linear regression* juga memberikan hasil yang sangat baik yang menandakan bahwa hubungan antar variabel pada data saham TLKM bersifat linier sehingga model sederhana seperti *linear regression* sudah cukup efektif untuk menangkap pola hubungan antarvariabel. Hasil tersebut menunjukkan perbedaan dengan model *neural network* (MLP), yang justru mengalami peningkatan performa setelah melalui proses *hyperparameter tuning*, khususnya pada konfigurasi dengan layer tersembunyi (*hidden_layer*) sebanyak [100,100] dan maksimum iterasi (*max_iter*) sebesar 300. Hasil tersebut menunjukkan bahwa meskipun *neural network* (MLP) mampu mengenali pola nonlinier yang lebih kompleks, model ini

memerlukan waktu pelatihan yang lebih lama dan sensitif terhadap parameter yang digunakan. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa *random forest* merupakan model paling optimal untuk memprediksi harga penutupan saham TLKM karena memiliki keseimbangan terbaik antara akurasi, stabilitas, dan efisiensi komputasi. Temuan ini sekaligus memperkuat bahwa pendekatan *machine learning* dapat menjadi metode yang efektif dalam analisis prediksi harga saham, terutama ketika dikombinasikan dengan pemilihan parameter yang tepat dan proses validasi yang menyeluruh.

Referensi

- [1] T. N. Bui, X. H. Nguyen, and K. T. Pham, "The Effect of Capital Structure on Firm Value: A Study of Companies Listed on the Vietnamese Stock Market," *Int. J. Financ. Stud.*, vol. 11, no. 3, 2023, <https://doi.org/10.3390/ijfs11030100>.
- [2] G. Karo, H. G. Barus, N. L. Salsabila, R. N. A. Harahap, R. Syahutra, and R. B. Rayoga, "Analyzing Factors Influencing Investment Decisions in the Indonesian Capital Market," *J. Ecoment Glob.*, vol. 10, no. 1, pp. 81–89, 2025, <https://ejournal.uigm.ac.id/index.php/EG/article/download/5383/2520/16517>.
- [3] E. S. Setiawan, "Pengaruh Debt to Equity Ratio, Earnings Per Share, dan Return On Equity terhadap Return Saham (Studi Pada Perusahaan Consumer Non-cyclical dan Consumer Cyclical yang Terdaftar di Bursa Efek Indonesia Tahun 2021-2022)," *Univ. Multimed. Nusant. Knowledge Cent.*, pp. 1–13, 2024, [Online]. Available: https://kc.umn.ac.id/id/eprint/34426/2/BAB_1.pdf
- [4] W. R. Murhadi, F. R. Kencanasar, and B. S. Sutedjo, "The Influence of Financial Literacy and Financial Interest on the Financial Risk Tolerance of Investor in Indonesia," *Journal of Law and Sustainable Development. Dev.*, vol. 11, no. 2, pp. 1–16, 2023, <https://ojs.journalsdg.org/jlss/article/download/310/250/878>.
- [5] I. D. Dichev and X. Zheng, "The volatility of stock investor returns," *J. Financ. Mark.*, vol. 70, p. 100927, 2024, <https://doi.org/10.1016/j.finmar.2024.100927>.
- [6] F. D. Sumaryana, N. Nugraha, M. Sari, and T. Heryawan, "The Influence of Macroeconomic Fundamentals and Investor Sentiment on the Indonesian Stock Exchange," *Atl. Press Int. BV*, no. Gebme 2023, pp. 7–13, 2024, doi: 10.2991/978-94-6463-443-3_2.
- [7] R. Rizky, P. Suwardi, E. Siska, and D. R. Sebastian, "Impact of Liquidity and Activity Ratios on Stock Price in Indonesia 's Telecom Companies," *J. Bus. Econ. Res.*, vol. 5, no. 3, pp. 495–503, 2024, <https://doi.org/10.47065/jbe.v5i3.5910>.
- [8] A. K. Dewa and S. Sundari, "Liquidity Amplifies, Dividends Stabilize: Evidence from Stock Price Volatility in Indonesia's LQ45 Index," *Formosa Journal of Multidisciplinary Research*, vol. 4, no. 8, pp. 3825–3840, 2025, <https://npaformosapublisher.org/index.php/fjmr/article/download/380/437>.
- [9] G. M. Tinungki, S. Siswanto, and A. Najiha, "The Gumbel Copula Method for Estimating Value at Risk: Evidence from Telecommunication Stocks in Indonesia during the COVID-19 Pandemic," *J. Risk Financ. Manag.*, vol. 16, no. 10, 2023, <https://doi.org/10.3390/rjfm16100424>.
- [10] B. Y. Dwiandiyanta, R. Hartanto, and R. Ferdiana, "Harnessing Deep Learning and Technical Indicators for Enhanced Stock Predictions of Blue-Chip Stocks on the Indonesia Stock Exchange (IDX)," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 20348–20357, 2025, <https://doi.org/10.48084/etasr.9850>.
- [11] Y. Huang, L. F. Capretz, and D. Ho, "Machine Learning for Stock Prediction Based on Fundamental Analysis," 2021 IEEE Symp. Ser. Comput. Intell. SSCI 2021 - Proc., 2021, <https://arxiv.org/pdf/2202.05702>.
- [12] J. Zou et al., "Stock Market Prediction via Deep Learning Techniques: A Survey," *arxiv.org*, vol. 1, no. 1, 2023, [Online]. Available: <https://arxiv.org/abs/2212.12717>.
- [13] M. Saberironaghi, J. Ren, and A. Saberironaghi, "Stock Market Prediction Using Machine Learning and Deep Learning Algorithms" *Sustain. Digit. Technol. Smart Cities Heal. Commun. Transp.*, pp. 127–138, 2023, <https://doi.org/10.1049/cit2.12059>.
- [14] Y. Li, H. Xue, S. Wei, R. Wang, and F. Liu, "A Machine Learning Approach for Investigating the Determinants of Stock Price Crash Risk: Exploiting Firm and CEO Characteristics," *Systems*, vol. 12, no. 5, 2024, <https://doi.org/10.3390/systems12050143>.
- [15] S. Ouf, M. El Hawary, A. Aboutabl, and S. Adel, "A Deep Learning-Based LSTM for Stock Price Prediction Using Twitter Sentiment Analysis," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 12, pp. 207–218, 2024, https://thesai.org/Downloads/Volume15No12/Paper_23-A_Deep_Learning_Based_LSTM_for_Stock_Price_Prediction.pdf.
- [16] T. Hall and K. Rasheed, "A Survey of Machine Learning Methods for Time Series Prediction," *Appl. Sci.*, vol. 15, no. 11, 2025, <https://www.mdpi.com/2076-3417/15/11/5957/pdf>.
- [17] X. Kong et al., "Deep learning for time series forecasting: a survey," *Int. J. Mach. Learn. Cybern.*, vol. 16, no. 7–8, pp. 5079–5112, 2025, <https://doi.org/10.48550/arXiv.2503.10198>.
- [18] W. Bao, Y. Cao, Y. Yang, H. Che, J. Huang, and S. Wen, "Data-driven stock forecasting models based on neural networks: A review," *Inf. Fusion*, vol. 113, p. 102616, 2025, <https://doi.org/10.1016/j.inffus.2024.102616>.
- [19] M. Barua, T. Kumar, K. Raj, and A. M. Roy, "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market," *FinTech*, vol. 3, no. 4, pp. 551–568, 2024, <https://doi.org/10.3390/fintech3040029>.
- [20] G. Campisi, S. Muzzioli, and B. De Baets, "A comparison of machine learning methods for predicting the direction of the US stock market on the basis of volatility indices," *International Journal of Forecasting*, vol. 40, no. 3, pp. 869–880, 2024, <https://doi.org/10.1016/j.ijforecast.2023.07.002>.
- [21] S. Abolmakarem, F. Abdi, K. Khalili-Damghani, and H. Didekhani, "A multi-stage machine learning approach for stock price prediction: Engineered and derivative indices," *Intell. Syst. with Appl.*, vol. 24, p. 200449, 2024, <https://doi.org/10.1016/j.iswa.2024.200449>.