

# Perbandingan Model *Random Forest* dan *Linear Regression* Untuk Memprediksi Harga Nikel

Deri Andika<sup>1\*</sup>, Rafli Yahya Praja Putra<sup>2</sup>, Pujianto<sup>3</sup>

<sup>1</sup>Informatika/Teknik dan Komputer

Universitas Baturaja

<sup>1\*</sup>deriandika253@gmail.com

<sup>2</sup>Informatika/Teknik dan Komputer

Universitas Baturaja

<sup>2</sup>raflibaturaja0@gmail.com

<sup>3</sup>Informatika/Teknik dan Komputer

Universitas Baturaja

<sup>3</sup>Pujianto.mail@email.com

**Abstrak**— Prediksi harga nikel menjadi perhatian penting dalam sektor industri dan investasi karena volatilitasnya yang tinggi dan perannya yang strategis dalam produksi baterai dan baja tahan karat. Penelitian ini bertujuan untuk membandingkan performa dua model *machine learning*, yaitu *Random Forest Regressor* dan *Linear Regression*, dalam memprediksi harga nikel menggunakan perangkat lunak Orange. Data historis harga nikel akan dianalisis dan dievaluasi berdasarkan metrik seperti *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, dan koefisien determinasi ( $R^2$ ). Diharapkan penelitian ini dapat mengidentifikasi model prediksi yang paling optimal dan memberikan kontribusi terhadap pengambilan keputusan strategis dalam *industry*.

**Kata kunci**— Data Mining, *Random Forest*, *Linear regression*, Orange, Harga Nikel, *Machine Learning*.

**Abstract**— *Nickel price prediction is a critical concern in industrial and investment sectors due to its high volatility and strategic role in battery and stainless steel production. This study aims to compare the performance of two machine learning models, namely Random Forest Regressor and Linear Regression, in predicting nickel prices using the Orange software. Historical nickel price data will be analyzed and evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination ( $R^2$ ). This study is expected to identify the most optimal predictive model and contribute to strategic decision-making in the metal commodity industry.*

**Keywords**— Data Mining, *Random Forest*, *Linear Regression*, Orange, Nickel Price, *Machine Learning*

## I. PENDAHULUAN

Nikel merupakan salah satu komoditas strategis dalam perekonomian di Indonesia maupun global, terutama karena perannya dalam produksi baja tahan karat dan baterai kendaraan listrik. Seiring dengan meningkatnya permintaan global, harga nikel menunjukkan volatilitas yang signifikan, dipengaruhi oleh berbagai faktor seperti kondisi geopolitik, dinamika pasar, dan kebijakan lingkungan. Oleh karena itu, prediksi harga nikel yang akurat menjadi krusial bagi pelaku industri, investor, dan pembuat kebijakan dalam merumuskan strategi yang efektif [4].

Dalam beberapa tahun terakhir, pendekatan *machine learning* telah digunakan secara luas dalam analisis data ekonomi dan prediksi harga komoditas. Dua metode yang sering digunakan adalah *Linear Regression* dan *Random Forest Regressor*. *Linear Regression* adalah metode

statistik klasik yang digunakan untuk memodelkan hubungan linier antara variabel independen dan dependen. Namun, metode ini memiliki keterbatasan dalam menangani hubungan non-linier dan kompleksitas data yang tinggi [8].

Sebaliknya, *Random Forest Regressor* adalah algoritma *ensemble learning* yang membangun sejumlah pohon keputusan dan menggabungkan hasilnya untuk meningkatkan akurasi prediksi. Metode ini lebih efektif dalam menangani data dengan hubungan non-linier dan dapat mengurangi risiko *overfitting*. Studi terbaru menunjukkan bahwa *Random Forest* sering kali memberikan hasil prediksi yang lebih akurat dibandingkan dengan *Linear Regression* dalam berbagai aplikasi, termasuk prediksi harga komoditas [5] [8].

Untuk memfasilitasi implementasi dan evaluasi model prediktif, digunakan perangkat lunak Orange, sebuah platform *open-source* berbasis *visual programming* yang mendukung berbagai algoritma *machine learning*, termasuk *Linear Regression* dan *Random Forest Regressor*[10]. Orange memungkinkan pengguna untuk membangun alur kerja analitik data secara intuitif dan efisien.

Penelitian ini bertujuan untuk membandingkan kinerja model *Random Forest Regressor* dan *Linear Regression* dalam memprediksi harga nikel menggunakan platform Orange. Hasil dari penelitian ini diharapkan dapat memberikan wawasan mengenai model prediktif yang lebih efektif untuk analisis harga nikel, serta memberikan kontribusi dalam pengambilan keputusan strategis di sektor industri dan investasi [8].

#### 1. Nikel

Nikel adalah logam transisi berwarna putih keperakan yang bersifat keras, ulet, dan tahan korosi. Dalam tabel periodik, nikel memiliki simbol Ni dan nomor atom 28. Nikel digunakan secara luas dalam industri, terutama untuk membuat paduan logam seperti stainless steel, serta digunakan dalam baterai, pelapisan logam, dan komponen elektronik [4].

#### 2. Data Mining

Data *mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan pembelajaran mesin (*machine learning*) mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* yang terkait [7].

#### 3. Software Orange

Orange merupakan software open source untuk pengolahan sebuah data *analytics* atau *data mining*. Dibandingkan dengan *software data mining* yang lainnya,

orange lebih unggul dalam hal visualisasi atau yang biasa disebut *visual programming*. Orange sudah menyediakan banyak widget yang telah diletakkan pada canvas atau *drawing board* yang kemudian dapat dihubungkan dengan widget lainnya. Melalui media canvas ini maka dapat memudahkan pemakai untuk bermain dengan data dan melakukan proses data analisis secara intuitif. Setiap widget yang ada memiliki fungsi masing-masing dan bisa menerima input dan mengeluarkan output [2].

#### 4. Random Forest

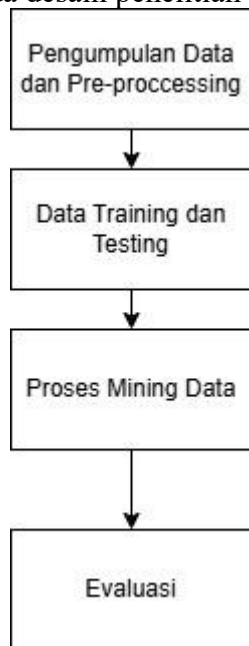
*Random Forest* merupakan algoritma dalam *machine learning* yang digunakan untuk pengklasifikasian dataset dalam jumlah besar. Karena fungsinya bisa dipakai dengan banyak dimensi dan dengan berbagai skala dan performa yang tinggi [12]. *Random Forest* diimplementasikan untuk berbagai sektor seperti perbankan, kesehatan, analisis keuangan, *e-commerce* dan sebagainya. Cara kerja algoritma ini adalah dengan membuat beberapa pohon keputusan dan menyartukannya untuk memperoleh prediksi yang lebih stabil serta akurat. Hutan (*Forest*) yang dibangun oleh algoritma tersebut merupakan kumpulan pohon keputusan dimana biasanya dilatih dengan metode *bagging*, yang diharapkan menghasilkan model *machine learning* yang lebih baik [5].

#### 5. Linear Regression

*Linear regression* merupakan salah satu metode statistik yang memberikan hasil output prediksi dengan melakukan pengembangan hubungan matematis antar variabel. Variabel independen merupakan variabel yang mempengaruhi sedangkan variabel dependen merupakan variabel yang dipengaruhi [3].

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental dengan metode komparatif. Tujuannya adalah membandingkan performa dua model prediksi *Random Forest Regressor* dan *Linear Regression* dalam memprediksi harga nikel [11]. Model dievaluasi berdasarkan tiga metrik akurasi: *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan koefisien determinasi ( $R^2$ ) [1]. Seluruh eksperimen dilakukan menggunakan perangkat lunak Orange Data Mining, yang mendukung pemrograman visual berbasis widget. Langkah-langkah yang diambil pada penelitian ini terdapat pada desain penelitian di bawah:



Gambar 1. Penelitian

1. Pengumpulan Data dan *Pre-Processing*

Data yang digunakan dalam penelitian ini adalah data historis harga nikel [7]. dalam penelitian ini di mulai dari awal tahun 2025, yaitu dari bulan januari sampai dengan april. Pengumpulan data dilakukan dengan mengunduh dataset dalam format CSV, Excel atau melalui API yang disediakan oleh penyedia data, atau bisa salin link berikut :

<https://databoks.katadata.co.id/tags/nikel>

2. Data *Training* dan *Testing*

*Dataset* dibagi menjadi dua bagian, yaitu training data (80%) dan testing data (20%) menggunakan widget (*Data Sampler*) [6].

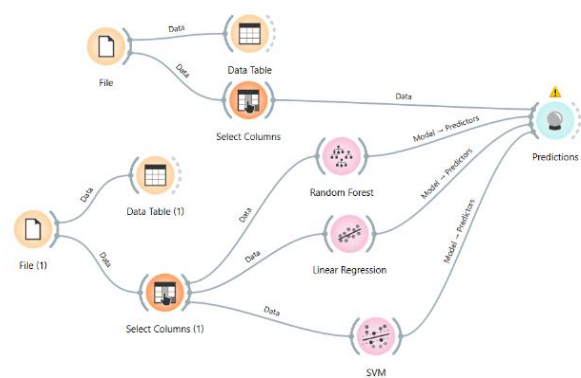
3. Proses *Mining* Data

Tahap ini melakukan penerapan teknik algoritma *Random Forest*, *Linear Regression*, dan SVM. Proses ini adalah tahapan yang menentukan model pohon keputusan yang sesuai dengan data training. Hasil penelitian dilakukan dengan menggunakan *tools* orange. Penggunaan data mining ini bertujuan untuk menentukan prediksi harga nikel [9][10].

4. Evaluasi

Tahap evaluasi dalam penelitian ini bertujuan untuk mengukur tingkat keakuratan model prediksi yang dibangun. Hasil evaluasi dari masing-masing model dibandingkan untuk menentukan model yang memiliki performa paling akurat dalam memprediksi harga nikel [12].

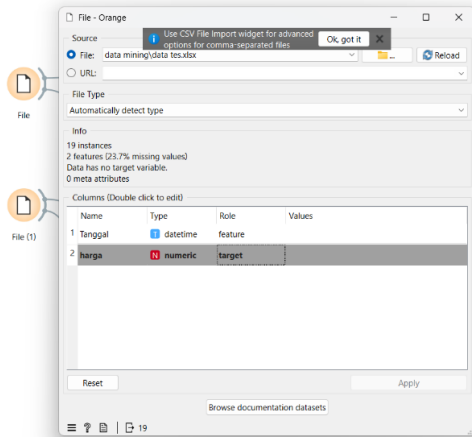
III. HASIL DAN PEMBAHASAN



Gambar 2. Model Penelitian

Hasil penelitian di buat berdasarkan dari tahapan awal penelitian sampai akhir penelitian dapat dilihat pada gambar 2.

A. *Widget File*



Gambar 3. Widget File

Pada gambar 3 menginput data harga nikel ke dalam *widget file* yang dimana datanya terdiri dari 19 *instance* dan 2 *feature*.

B. *Widget Data Tabel Testing*

1. *Data Tabel Testing*

	harga	Tanggal
1	15860	2025-04-03
2	15295	2025-04-04
3	14550	2025-04-07
4	14510	2025-04-08
5	14030	2025-04-09
6	14845	2025-04-10
7	15220	2025-04-11
8	15385	2025-04-14
9	15590	2025-04-15
10	15605	2025-04-16
11	?	2025-04-17
12	?	2025-04-18
13	?	2025-04-19
14	?	2025-04-20
15	?	2025-04-21
16	?	2025-04-22
17	?	2025-04-23
18	?	2025-04-24
19	?	2025-04-25

Gambar 4. Data Tabel Testing

Data table *testing* tempat untuk melihat data yang sudah di input kedalam data *file*, Gambar 4 adalah bentuk data yang digunakan dalam penelitian.

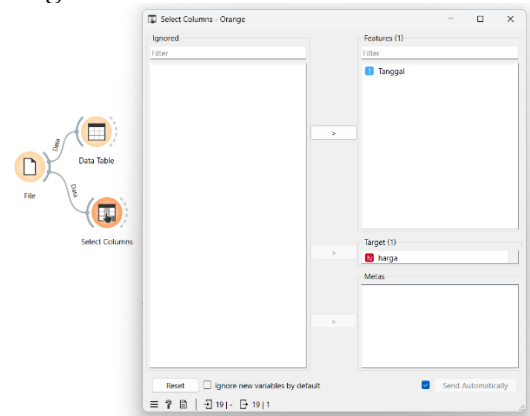
2. *Data Tabel Training*

	Harga	Tanggal
29	15350	2025-02-26
30	15530	2025-02-26
31	15820	2025-02-27
32	15675	2025-02-28
33	15770	2025-03-03
34	15970	2025-03-04
35	15975	2025-03-05
36	16180	2025-03-06
37	16270	2025-03-07
38	16650	2025-03-10
39	16535	2025-03-11
40	16710	2025-03-12
41	16455	2025-03-13
42	16655	2025-03-14
43	16530	2025-03-17
44	16300	2025-03-18
45	16575	2025-03-19
46	16210	2025-03-20
47	16145	2025-03-21
48	16105	2025-03-24
49	16050	2025-03-25
50	16200	2025-03-26
51	16275	2025-03-27
52	16375	2025-03-28
53	15975	2025-03-31
54	16060	2025-04-01
55	16020	2025-04-02
56	15860	2025-04-03
57	15295	2025-04-04
58	14550	2025-04-07
59	14510	2025-04-08
60	14030	2025-04-09
61	14845	2025-04-10
62	15220	2025-04-11
63	15385	2025-04-14
64	15590	2025-04-15
65	15605	2025-04-16

Gambar 5. Data Tabel Training

Data *table training* tempat untuk melihat data yang sudah di input kedalam data *file*, Gambar 5 adalah bentuk data yang digunakan dalam penelitian.

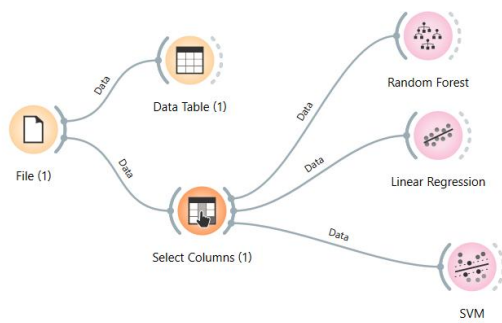
C. *Widget Select Columns*



Gambar 6. Select Column

Pada gambar 6 setelah data terhubung ke *select columns* ambil harga dan di jadikan sebagai target *variable*, selanjutnya akan dilakukan pengujian terhadap data tersebut menggunakan algoritma *Random Forest* dan *Linear Regression*, kami juga menambahkan metode SVM sebagai tambahan untuk pembandingan.

D. Penerapan Algoritma Random Forest, Linear Regression, dan SVM



Gambar 7. Penerapan Model Algoritma

Pada gambar 7 menambahkan model algoritma dari *select columns* ke model *random forest*, *linear regression*, dan SVM. Selanjutnya akan di hubungkan ke prediction dan data pada *select column (data testing)* juga di hubungkan untuk melihat hasil dari prediksi harga nikel.

E. Hasil Predictions

Predictions - Orange								
Shown regression error: Difference								
	Random Forest	error	Linear Regression	error	SVM	error	harga	Tanggal
1	15.948	-1584...	15.753	-1584...	15.760	-1584...	15860	2025-04-03
2	15.134	-1527...	15.754	-1527...	15.395	-1527...	15295	2025-04-04
3	14.567	-1453...	15.755	-1453...	14.756	-1453...	14550	2025-04-07
4	14.567	-1449...	15.756	-1449...	14.756	-1449...	14510	2025-04-08
5	14.567	-1401...	15.756	-1401...	14.756	-1401...	14030	2025-04-09
6	14.655	-1483...	15.757	-1482...	14.945	-1483...	14845	2025-04-10
7	15.039	-1520...	15.757	-1520...	15.320	-1520...	15220	2025-04-11
8	15.441	-1536...	15.759	-1536...	15.485	-1536...	15385	2025-04-14
9	15.441	-1557...	15.759	-1557...	15.690	-1557...	15590	2025-04-15
10	15.441	-1558...	15.760	-1558...	15.705	-1558...	15605	2025-04-16
11	15.441		15.760		15.756		?	2025-04-17
12	15.441		15.761		15.756		?	2025-04-18
13	15.441		15.761		15.756		?	2025-04-19
14	15.441		15.762		15.756		?	2025-04-20
15	15.441		15.762		15.756		?	2025-04-21
16	15.441		15.763		15.756		?	2025-04-22
17	15.441		15.763		15.756		?	2025-04-23
18	15.441		15.764		15.756		?	2025-04-24
19	15.441		15.764		15.756		?	2025-04-25

Gambar 8. Hasil Prediction

Gambar 8 menunjukkan output dari widget Predictions di Orange Data Mining, yang membandingkan hasil prediksi tiga model regresi: *Random Forest*, *Linear Regression*, dan SVM (*Support Vector Machine*).

F. Hasil Perhitungan MSE, RMSE, MAE, MAPE, R2

Model	MSE	RMSE	MAE	MAPE	R2
Random Forest	284016.478	532.932	408.216	0.027	0.074
Linear Regression	459747.748	678.047	542.822	0.036	-0.500
SVM	376050.000	613.229	519.500	0.035	-0.227

Gambar 9. Hasil MSE, RMSE, MAE, MAPE, R2

IV. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan terhadap model prediksi harga nikel menggunakan algoritma *Random Forest*, *Linear Regression*, dan SVM melalui perangkat lunak Orange, diperoleh kesimpulan sebagai berikut:

- Model *Random Forest* memberikan performa terbaik dibandingkan dua model lainnya, dengan nilai MSE sebesar 284016.48, RMSE 532.93, MAE 408.22, dan MAPE 0.027, serta satu-satunya model dengan nilai R<sup>2</sup> positif sebesar 0.074. Hal ini menunjukkan bahwa *Random Forest* mampu menangkap pola data dengan lebih baik dan memberikan prediksi yang lebih akurat.
- Linear Regression* dan SVM menunjukkan hasil prediksi yang kurang memuaskan dengan nilai R<sup>2</sup> negatif (-0.500 untuk *Linear Regression* dan -0.227 untuk SVM), yang mengindikasikan bahwa kedua model tidak dapat menjelaskan variabilitas data dengan baik. Selain itu, nilai error pada kedua model juga lebih tinggi dibandingkan *Random Forest*.
- Dari hasil tabel prediksi, model *Random Forest* menunjukkan hasil prediksi yang lebih variatif dan mengikuti pola historis, sementara *Linear Regression* dan SVM menghasilkan prediksi yang cenderung stagnan pada rentang data tanpa nilai aktual (tanggal 17–25 April 2025).
- Berdasarkan evaluasi keseluruhan, model *Random Forest* direkomendasikan sebagai algoritma terbaik untuk digunakan dalam memprediksi harga nikel, khususnya pada data yang bersifat fluktuatif dan non-linier seperti harga komoditas.

Penelitian ini masih memiliki keterbatasan pada jumlah data historis yang digunakan dan belum mempertimbangkan faktor eksternal seperti harga komoditas lain, nilai tukar, atau kondisi pasar global. Pengembangan selanjutnya dapat dilakukan dengan menambahkan fitur eksternal serta mengeksplorasi model prediktif lainnya seperti XGBoost, LSTM, atau model *deep learning* lainnya.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada dosen pembimbing yang telah memberikan arahan, masukan, dan semangat selama proses penelitian ini berlangsung. Terima kasih juga kepada seluruh teman dan rekan seperjuangan yang senantiasa memberi dukungan, baik secara akademik maupun emosional. Tidak lupa, penulis juga menghaturkan rasa terima kasih kepada pihak-pihak yang telah menyediakan data, referensi, dan fasilitas yang sangat membantu kelancaran penelitian ini. Semoga segala bentuk bantuan yang diberikan menjadi amal kebaikan yang tak ternilai. Akhir kata, semoga karya sederhana ini dapat bermanfaat dan menjadi bagian kecil dari kontribusi dalam pengembangan ilmu pengetahuan di bidang data mining dan prediksi harga komoditas.

#### REFERENSI

- [1] Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes. *Netw. Eng. Res. Oper.*, 6(2), 133.
- [2] Indriyanti, I., Ichsan, N., Fatah, H., Wahyuni, T., & Ermawati, E. (2022). Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin. *Jurnal Responsif: Riset Sains dan Informatika*, 4(2), 118-125.
- [3] Duran, P. A., Vitianingsih, A. V., Riza, M. S., Maukar, A. L., & Wati, S. F. A. (2024). Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression. *Teknika*, 13(1), 27-34.
- [4] resnu Maulana, A. (2024). PERBANDINGAN METODE FTS LEE DAN FTS SAXENA EASO PADA PREDIKSI HARGA NIKEL. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(2).
- [5] Khalim, K. A., Hayati, U., & Bahtiar, A. (2023). Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 498-504.
- [6] Muharrom, M. (2023). Analisis Penggunaan Orange Data Mining untuk Prediksi Harga USDT/BIDR Binance. *Bulletin of Information Technology (BIT)*, 4(2), 178-184.
- [7] Ramdhani, F., & Setiawan, K. (2024). Penerapan Data Mining untuk Prediksi Pelanggan di PT. XYZ Menggunakan Algoritma Linear Regression: Application of Data Mining for Customer Prediction at PT. XYZ Using Linear Regression Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 490-497.
- [8] Fitri, E. (2023). Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah. *Journal of Applied Computer Science and Technology*, 4(1), 58-64.
- [9] Surbakti, N. K. (2021). Data Mining Pengelompokan Pasien Rawat Inap Peserta BPJS Menggunakan Metode Clustering (Studi Kasus: RSUD. Bangkatan). *Journal of Information and Technology*, 1(2), 47-53.
- [10] Mahena, Y., Rusli, M., & Winarso, E. (2015). Prediksi Harga Emas Dunia Sebagai Pendukung Keputusan Investasi Saham Emas Menggunakan Teknik Data Mining. *Kalbiscentia J. Sains dan Teknol.*, 2(1), 36-51.
- [11] Handoko, A. C., & Hendry, H. (2023). Perbandingan Metode Supervised Learning Untuk Prediksi Diabetes Gestasional Dengan Software Orange. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 8(4), 1238-1247.
- [12] Pranadjaya, E., Pangestu, E. S., Sereati, C. O., Octaviani, S., & Darmawan, M. (2024). Perbandingan Algoritma Machine Learning menggunakan Orange Data Mining untuk Klasifikasi Jenis Kendaraan pada Sistem Tilang Digital. *Jurnal Elektro*, 17(1), 41-47.