

Analisis Komparasi Algoritma Data Mining Naive Bayes, K-Nearest Neighbors dan Regresi Linier Dalam Prediksi Harga Emas

Muhammad Muharrom

Fakultas Teknik dan Informatika, Program Studi Teknologi Informasi, Universitas Bina Sarana Informatika, Jakarta Pusat, Indonesia

Email: muhammad.muu@bsi.ac.id

(* : coressponding author: muhammad.muu@bsi.ac.id)

Abstrak– Hasil implementasi Orange data Mining untuk forecasting nilai Harga Emas ditampilkan pada widget Test and Score. Didapatkan nilai RMSE dan MAE dari masing-masing model dari pengujian tersebut. Nilai RMSE dan MAE untuk metode K-Nearest Neighbor (K-NN) masing-masing adalah 0.007 dan 0.006, sedangkan untuk metode Support Vector Machine (SVM) adalah 0.006 dan 0.005. Nilai RMSE dan MAE untuk metode Regresi Linear masing-masing 0.004 dan 0.003. Dibandingkan dengan metode K-Nearest Neighbor dan SVM, metode Regresi Linier paling baik dalam memprediksi perubahan harga Emas berdasarkan data RMSE dan MAE tersebut di atas. Untuk penelitian mendatang, metode praktik terbaik ini perlu dikaji lebih dalam. Disarankan kepada penelitian selanjutnya untuk membandingkan metode Regresi Linier dengan pendekatan alternatif yang menggunakan tool set Orange atau tools terkait lainnya.

Kata Kunci: Orange; Data Mining; Emas; Regresi Linier; Prediksi;

Abstract– *The results of implementing Orange Data Mining for forecasting the value of the Gold Price are displayed on the Test and Score widget. RMSE and MAE values were obtained from each model from the test. The RMSE and MAE values for the K-Nearest Neighbor (K-NN) method are 0.007 and 0.006, respectively, while for the Support Vector Machine (SVM) method are 0.006 and 0.005. The RMSE and MAE values for the Linear Regression method are 0.004 and 0.003, respectively. Compared to the K-Nearest Neighbor and SVM methods, the Linear Regression method is the best at predicting changes in Gold prices based on the RMSE and MAE data mentioned above. For future research, this best practice method needs to be studied more deeply. It is recommended for future research to compare the Linear Regression method with alternative approaches using the Orange tool set or other related tools.*

Keyword: Orange; Data Mining; Gold; Linier Regression; Prediction;

1. PENDAHULUAN

Investasi adalah strategi yang dilakukan dengan harapan mendapatkan keuntungan finansial atau pertumbuhan dimasa depan. Jenis investasi yang paling mendasar adalah melalui penggunaan emas[1]. Selain dapat tumbuh dalam waktu yang lama, logam mulia atau emas juga memiliki karakteristik yang sulit berkarat bahkan setelah beberapa tahun, karena memiliki urat yang lunak, tahan korosi, dan mudah terurai. Pemasukan emas terbatas dan tidak mudah ditawarkan, sedangkan permintaan emas tidak pernah menurun. Hal ini terkait dengan fakta bahwa harga emas atau logam mulia lainnya semakin stabil dan turun serta semakin mengalami inflasi. Emas atau logam mulia menjadi jenis investasi yang sering disebut bebas risiko jika dibandingkan dengan jenis investasi lainnya. Dikatakan bahwa berinvestasi saham itu sederhana karena tidak harus dimiliki oleh orang kaya atau individu dengan keahlian khusus. Emas seperti ini juga bisa digunakan oleh siapa saja dari golongan manapun. Berinvestasi emas pada umumnya bukanlah keputusan yang bijak, namun merupakan keputusan yang sangat menguntungkan karena merupakan satu-satunya jenis log mulia yang memiliki harga yang tidak berubah ketika terjadi inflasi. Ketahanannya pada saat inflasi adalah satu-satunya penjelasan mengapa emas menjadi pilihan yang lebih disukai untuk melakukan investasi karena zaman emas telah berubah dari waktu ke waktu.[2] Penerapan aplikasi orange data mining dalam memprediksi harga Emas dengan metode K-Nearest Neighbor (K-NN), Support Vector Machine (SVM) dan Regresi Linier (Linier Regression). Prediksi kinerja dari ketiga metode yang digunakan dapat dilihat dari hasil Root Mean Square Error (RMSE) dan Mean Absolute Error (MAE). Dataset yang telah dikumpulkan akan dianalisa menggunakan aplikasi orange data mining, proses pengujian menerapkan K-Fold Cross Validation (K=5), sedangkan proses perbandingan metode menggunakan Test and Score. [3]

Dalam data mining fungsi merupakan salah satu bagian dari fungsi minor. Prediksi adalah cara untuk mencari kemungkinan hasil pada masa akan datang. Pada teknologi data mining, berdasarkan bukti dan temuan pada data yang digunakan metode peramalan dapat memprediksi kemungkinan pada masa datang. Perhitungan matematik dan statistik berperan erat dalam proses prediksi. [4]

Tools yang biasanya digunakan untuk menilai kinerja suatu metode antara lain WEKA dan Rapid Miner. Kinerja suatu metode untuk melakukan prediksi dapat dilakukan juga dengan menggunakan orange tools. Orange data mining merupakan aplikasi open source yang mampu membantu penelitian dalam menganalisa suatu data [5]

Penelitian ini bertujuan untuk mendapatkan metode terbaik untuk memprediksi harga emas. Pada penelitian ini akan menerapkan orange tools untuk memprediksi harga emas dengan menggunakan metode Linear Regresi, SVM dan k-NN. Kedua metode tersebut akan di evaluasi dan dibandingkan kinerjanya sehingga dapat diketahui metode yang paling baik untuk memprediksi harga Emas berdasarkan nilai RMSE dan nilai MAE dari masing-masing metode. Nilai tersebut dapat dilihat menggunakan widget Test and Score pada aplikasi Orange.

Penelitian sebelumnya terkait dengan prediksi emas yaitu: data set yang digunakan selama 3 tahun yaitu dari Januari 2016 – Desember 2019 dan data training prediksi 1 tahun kedepan dengan asumsi harga stabil, dengan pengujian sebanyak 5 kali 52 data, di mana dalam penggunaan dalam melihat peluang berdasarkan histori kejadian yang terjadi sehingga akurasi cukup baik sebesar 95,92% dalam penggunaan serta melihat peluang profit ya atau tidaknya harga emas[6].

Menggunakan metode algoritma k-nn dapat diterapkan pada berbagai objek dataset lainnya seperti menggunakan analisis sentimen film untuk mengkategorikan data sebagai negatif atau positif serta menentukan mana yang paling akurat. Hasil pengujian komparatif kedua algoritma menunjukkan bahwa algoritma Nave Bayes memiliki akurasi rata-rata yang lebih tinggi yaitu 99,63% dengan AUC sekitar 1.000, sedangkan algoritme K-NN memiliki akurasi rata-rata yang lebih tinggi yaitu 99,25% dengan AUC sebesar 1.000 [7].

Penelitian terkait dengan metode linier regresi yaitu pada prediksi cryptocurrency, hasilnya adalah model regresi linier yang dapat digunakan untuk memprediksi harga cryptocurrency pada periode selanjutnya. Hasil dari penelitian ini adalah model linier regresi yang dapat memprediksi harga cryptocurrency dengan akurasi sebesar 91,46%. Dengan demikian, metode linier regresi dapat digunakan sebagai alternatif untuk memprediksi harga cryptocurrency[8].

Penelitian terkait lainnya menggunakan regresi linier yaitu analisis metode Single Moving Averaga dalam meramalkan kurs tengah BI dolar ke Rupiah, didapatkan rata-rata MAPE sebesar 0,19080% dengan status bobot “sangat baik” dan terjadi pengurangan sebesar 0,00180% dibandingkan hasil analisis metode Regresi Linier Sederhana. Namun, selisih yang dihasilkan masih tetap tergolong besar yaitu berupa ribuan Rupiah. Sebagai contoh, pada tanggal 24 Juni 2022 harga beli emas digital pegadaian berada di angka Rp.918000 dan hasil ramalan harga beli emas digital pegadaian berada di Rp.916753, dengan selisih Rp.1247. Oleh karena itu, peneliti mengambil kesimpulan bahwa kombinasi metode Regresi Linier Sederhana dan Single Moving Average bernilai baik dari penilaian MAPE namun masih kurang baik dalam meramalkan harga beli emas digital pegadaian.[9].

Penggunaan regresi linier juga dapat memprediksi IHSG berdasarkan faktor-faktor yang mempengaruhi berdasarkan penelitian sebelumnya yaitu harga minyak mentah, harga emas dan kurs. Setelah ditambah variabel moderasi faktor yang mempengaruhi IHSG bertambah dua variabel yaitu interaksi antara variabel harga emas dan kurs serta variabel harga minyak mentah dan kurs. Tetapi kursnya menjadi tidak berpengaruh terhadap IHSG. Pada kasus ini disebut dengan Pure Moderasi yaitu variabel moderasi yang memoderasi hubungan antara variabel independen dan variabel dependen dimana variabel Z berinteraksi dengan variabel independen tanpa menjadi variabel independen.[10]

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Untuk mempermudah dalam pengerjaan penelitian ini, maka penulis membuat tahapan penelitian seperti pada gambar 1 berikut ini.



Gambar 1. Tahapan Penelitian

Tahapan pada penelitian ini terdiri dari 4 tahap yaitu pengumpulan dataset dan preprocessing, pembuatan data latih dan data uji, proses prediksi harga emas, evaluasi kinerja dan hasil perbandingan metode.

1. Pengumpulan Dataset dan *Preprocessing*
 Dataset yang digunakan adalah data Emas dengan rentang waktu 06 Juni 2023 sampai dengan 06 Juli 2023 dari laman <https://id.investing.com/commodities/gold-historical-data>.
 Pembuatan data latih dan uji. Atribut yang digunakan pada penelitian ini antara lain:
 - a. Tanggal: Tanggal Transaksi Emas
 - b. Harga: harga dalam Emas pada tanggal yang terkait.
 - c. Pembukaan: nilai tukar Emas pada hari yang terkait.
 - d. Harga tertinggi Emas pada hari yang terkait.
 - e. Terendah: harga terendah Emas pada hari hari yang terkait.
 - f. Volume: total transaksi perdagangan Emas pada tanggal terkait.
 - g. Perubahan% : perubahan kurs harga Emas pada tanggal terkait terhadap harga Emas pada tanggal sebelumnya.
2. Pembuatan sampling data Uji dan Testing
 Sampling dataset yang digunakan dapat dilihat pada Tabel 1 di bawah ini.

Tabel 1. Sampling Dataset

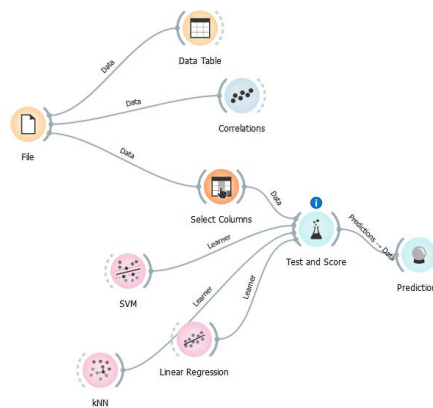
Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%
06/07/2023	1,916.60	1,923.40	1,933.95	1,908.55	0	-0.54%

05/07/2023	1,927.10	1,929.10	1,942.90	1,922.10	245,93K	-0.35%
04/07/2023	1,933.95	1,928.85	1,938.05	1,927.75	0	0.23%
03/07/2023	1,929.50	1,927.80	1,939.90	1,917.70	152,96K	0.01%
30/06/2023	1,929.40	1,916.40	1,930.80	1,908.10	180,69K	0.60%
29/06/2023	1,917.90	1,916.50	1,921.30	1,900.60	205,25K	0.24%
28/06/2023	1,913.40	1,915.20	1,916.70	1,903.60	0,63K	-0.09%
27/06/2023	1,915.10	1,927.40	1,930.20	1,912.30	0,23K	-0.52%
26/06/2023	1,925.10	1,926.40	1,934.50	1,923.00	0,77K	0.21%
23/06/2023	1,921.00	1,915.00	1,939.30	1,911.30	0,75K	0.31%
22/06/2023	1,915.00	1,935.90	1,935.90	1,914.00	0,37K	-1.09%
21/06/2023	1,936.10	1,939.50	1,940.70	1,922.20	0,81K	-0.14%
20/06/2023	1,938.90	1,961.70	1,961.70	1,932.40	0,74K	-1.28%
19/06/2023	1,964.05	1,969.55	1,970.85	1,959.90	0	-0.27%
18/06/2023	1,969.45	1,970.85	1,971.70	1,968.65	0	0.37%
16/06/2023	1,962.20	1,961.30	1,970.60	1,956.70	0,62K	0.02%
15/06/2023	1,961.80	1,948.80	1,963.70	1,929.60	0,82K	-0.36%
14/06/2023	1,968.90	1,957.20	1,973.90	1,952.50	199,71K	0.53%
13/06/2023	1,958.60	1,971.40	1,985.90	1,953.00	198,40K	-0.56%
12/06/2023	1,969.70	1,975.60	1,981.40	1,963.10	122,54K	-0.38%
09/06/2023	1,977.20	1,980.50	1,987.80	1,971.00	130,41K	-0.07%
08/06/2023	1,978.60	1,956.20	1,985.70	1,955.20	191,84K	1.03%
07/06/2023	1,958.40	1,980.00	1,986.50	1,955.40	189,53K	-1.17%
06/06/2023	1,981.50	1,978.00	1,982.90	1,970.30	136,00K	0.36%

Pada tahap ini dilakukan pengubahan value tanggal yang bertipe string menjadi numeric dan karena pada data ini ada yang kosong (*missing value*) maka diberikan nilai 0 sehingga dapat diproses dalam perhitungan algoritma.

3. Proses Data Mining Prediksi Harga Emas

Model prediksi harga emas dapat dianalisa menggunakan orange tool untuk memilih metode terbaik. Model prediksi dataset harga bitcoin dapat dilihat pada Gambar 2. dibawah ini.



Gambar 2. Desain model Orange Data mining

4. Evaluasi Kinerja dan hasil perbandingan metode

Setelah dilakukan preprocessing data, maka selanjutnya dataset diproses kedalam model prediksi pada *software* orange dengan menggunakan metode K-NN dan SVM dan Linear Regresi.

Tujuan dari penelitian ini untuk melakukan analisa perbandingan metode K- NN, SVM Regresi Linier untuk memprediksi harga emas dengan menggunakan Orange Tools untuk simulasinya.

2.2 K-Nearest Neighbors

Konsep dasar dari K-NN adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga terdekatnya dalam data pelatihan. [11] K-Nearest Neighbors (K-NN) merupakan algoritma pembelajaran mesin non-parametrik dan berbimbing yang digunakan untuk tugas klasifikasi maupun regresi. Algoritma ini didasarkan pada konsep bahwa data yang serupa cenderung memiliki hasil yang serupa pula. K-NN bekerja dengan menghitung jarak antara data baru dengan data tetangga terdekatnya dalam ruang fitur. Untuk tugas regresi, K-NN memprediksi nilai dari data baru dengan cara mengambil rata-rata nilai dari k tetangga terdekatnya. Nilai yang diprediksi adalah rata-rata (atau median) dari variabel target dari data tetangga. K-NN tidak membuat asumsi kuat tentang distribusi data yang mendasarinya, sehingga algoritma ini sangat fleksibel. Namun, K-NN memiliki kompleksitas komputasi yang tinggi, terutama untuk dataset yang besar, karena membutuhkan perhitungan jarak antara setiap pasang data. Selain itu, performa K-NN dapat sensitif terhadap pemilihan metrik jarak dan nilai k. Untuk mengevaluasi kinerja algoritma K-NN, metrik seperti RMSE (Root Mean Squared Error) dan MAE (Mean Absolute Error) dapat digunakan. Metrik ini mengukur perbedaan antara nilai prediksi dan nilai aktual. Nilai RMSE dan MAE yang lebih rendah menunjukkan performa prediksi yang lebih baik. Merupakan suatu metode yang menggunakan algoritma supervised dimana hasil dari Query instance dapat diklasifikasikan berdasarkan mayoritas dari label class pada K-NN[1]. Algoritma K-NN bekerja berdasarkan jarak terpendek dari Query instance ke training data untuk menentukan K-NNnya dan merupakan salah satu cara untuk menghitung jarak dekat [12]

2.3. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah pengklasifikasi diskriminatif yang menghasilkan hyperplane pemisah. Toleransi kesalahan disertakan untuk membuat hyperplane pemisah menjadi kuat jika terjadi data kelas yang tidak dapat dipisahkan. Dalam Algoritma SVM ada trik kernel dimana ada SVM linear dan SVM nonlinear. SVM hyperplane linear bekerja hanya pada data yang dapat dipisahkan dengan cara linear. SVM Non Linear yaitu data yang berdistribusi pada kelas yang tidak linear sering digunakan pendekatan kernel pada fitur awal set. [13] SVM merupakan suatu teknik yang digunakan untuk melakukan prediksi pada kasus klasifikasi maupun regresi. Konsep klasifikasi dengan SVM adalah mencari hyperlane terbaik yang memiliki fungsi sebagai pemisah dari dua kelas data, opini positif (+1) dan opini negatif(-1). [14]

2.4 Regresi Linier (*Linear Regression*)

Regresi linier adalah teknik pemodelan statistik yang digunakan untuk menganalisis hubungan antara variabel dependen dan satu atau lebih variabel independen. Ini adalah metode yang banyak digunakan dalam analisis data dan tugas prediksi. Tujuan regresi linier adalah menemukan persamaan linier terbaik yang menggambarkan hubungan antara variabel. Dalam konteks memprediksi harga emas, regresi linier dapat digunakan untuk memodelkan hubungan antara variabel independen (seperti waktu, tren pasar, atau faktor lain yang relevan) dan variabel dependen (harga emas). Model regresi linier bertujuan untuk memperkirakan koefisien persamaan, yang mewakili bobot atau kontribusi variabel independen dalam memprediksi variabel dependen. Ada 2 tipe variabel dalam regresi linier yaitu: variabel pemberi pengaruh dan variabel terpengaruh. Variabel pemberi pengaruh digunakan sebagai sebab, sedangkan variabel terpengaruh sebagai akibat. penggunaan model regresi linear sederhana pada data historis harga dapat digunakan untuk memprediksi pergerakan harga pada proses penelitiannya.[15] Regresi Linier merupakan metode statistika yang menjelaskan tentang pola model untuk membentuk hubungan antara variabel terikat dan variabel bebas atau lebih. Tujuan dari metode ini adalah untuk memprediksi nilai y untuk nilai x yang diberikan. Model regresi linier sederhana adalah model regresi yang paling sederhana[16]

2.5 Root Mean Square Error (RMSE)

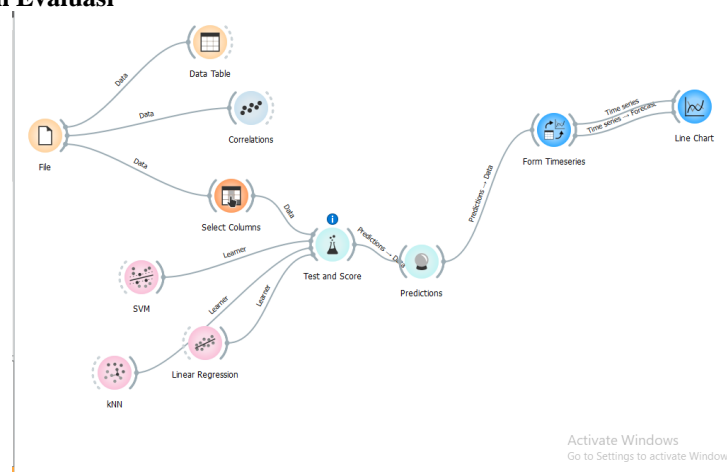
Keakuratan dari sebuah peramalan diperlukan evaluasi terhadap data yang sebenarnya. Ada beberapa metode yang sering digunakan untuk melakukan perhitungan kesalahan peramalan, salah satunya adalah root mean square error (RMSE). Metode ini sering digunakan dalam mengevaluasi hasil peramalan. Ukuran yang sering digunakan dari perbedaan antara nilai- nilai diprediksi oleh model atau estimator dan nilai-nilai yang benar-benar diamati yaitu RMSE. RMSE adalah salah satu teknik umum yang dapat dipakai untuk mengetahui tingkat error pada model prediksi data yang berupa data angka. Hasil dari RMSE didapatkan dari rata-rata kuadrat total error yang didapatkan pada model prediksi.[17] Semakin kecil nilai RMSE, semakin baik tingkat akurasi prediksinya. Keakuratan sebuah model peramalan dalam melakukan prediksi ditentukan oleh nilai terkecil dari masing-masing metode akurasi data.[18]

2.6 Mean Absolute Error (MAE)

Mean Absolute Error (MAE) adalah salah satu metode yang digunakan untuk mengukur tingkat keakuratan model peramalan. Hasil Mean Absolute Error menunjukkan nilai rata-rata kesalahan mutlak atau absolut dari nilai sebenarnya dengan nilai peramalan. Keakuratan dari hasil perhitungan prediksi menggunakan metode Mean Absolute Error (MAE) yang digunakan untuk mengukur tingkat keakuratan suatu mode peramalan. rata-rata selisih mutlak nilai sebenarnya (aktual) dengan nilai perkiraan (prediksi) [19]. MAE merupakan ilai rata-rata dari selisih absolut antara nilai sebenarnya dan nilai yang diprediksi oleh model. Nilai MAE yang lebih kecil menunjukkan bahwa model tersebut lebih akurat dalam memprediksi nilai yang sebenarnya terjadi.[20]

3. HASIL DAN PEMBAHASAN

3.1 Proses Pembahasan Evaluasi

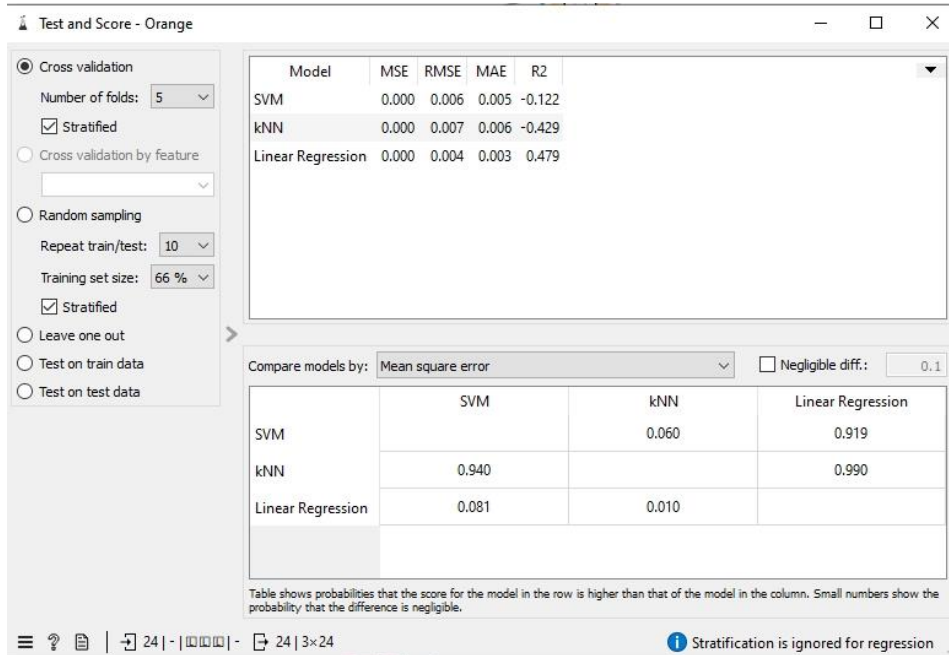


Gambar 3. Design untuk menghitung keberhasilan model prediksi. Sumber: Peneliti (2023)

Pada Gambar 3. Merupakan proses evaluasi hasil perbandingan dari model prediksi yang diuji. Proses perhitungan keberhasilan model prediksi pada orange tool dapat menggunakan *widjet Test and Score* dan hasil prediksinya dapat dilihat menggunakan *widjet Predictions* dan *line chart* idget dari *timeseries* untuk melihat pemodelan bentuk diagram.

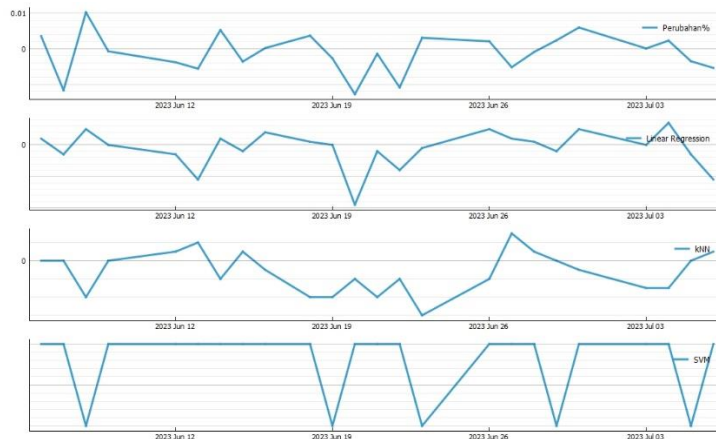
1. Proses Simulasi Hasil Prediksi

Data yang telah melalui tahap preprocessing selanjutnya dilakukan pengujian untuk mendapatkan model prediksi terbaik. Model prediksi yang telah dilakukan pengujian dan evaluasi dengan menggunakan kumpulan data uji pada aplikasi orange dimana 1 atribut sebagai targetnya didapatkan hasil simulasi seperti pada Gambar 4.



Gambar 4. Hasil Test and Score

Gambar 4 merupakan Aplikasi widget Tes dan Skor dari aplikasi Orange. Pada penelitian ini proses pengujian menggunakan K-Fold Cross Validation (K=5) yang dapat diatur pada widget Test and Score seperti yang dilihat pada gambar. Pada widget yang sama juga ditampilkan hasil dari kedua metode evaluasi tersebut, dimana hasil perbandingan dari kedua model prediksi harga Emas ditunjukkan dengan nilai MSE, RMSE, MAE, dan R2. Analisis saat ini hanya mempertimbangkan koefisien MAE dan RMSE untuk memeriksa kinerja model prediksi yang dianalisa. Hasil evaluasi menunjukkan bahwa kedua model prediksi harga Emas memiliki tingkat akurasi yang berbeda. Support Vector Machine (SVM) memiliki koefisien regresi yang lebih rendah untuk MAE linier dan RMSE dibandingkan dengan K-NN. Hal ini menunjukkan bahwa metode terbaik untuk memperkirakan harga Emas yaitu regresi linier.



Gambar 5. Grafik Perubahan Harga Hasil Prediksi keseluruhan metode.

Grafik pada gambar 5 menunjukkan hasil dari ketiga metode yang digunakan untuk memprediksi perubahan nilai harga Emas. Hasil prediksi harga menggunakan metode Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), dan Regresi Linear. Hasil menunjukkan bahwa metode Regression Linear jauh lebih baik dibandingkan dengan

metode SVM dan K-NN. Hal ini dikarenakan metode K-NN memiliki tingkat error yang sangat tinggi dan secara konsisten menghasilkan prediksi yang tidak akurat terhadap perubahan harga Emas. Hasil prediksi dengan metode Regresi Linear sejalan dengan prediksi harga Emas dengan margin error yang kecil.

2. Hasil Prediksi Metode Regresi Linier

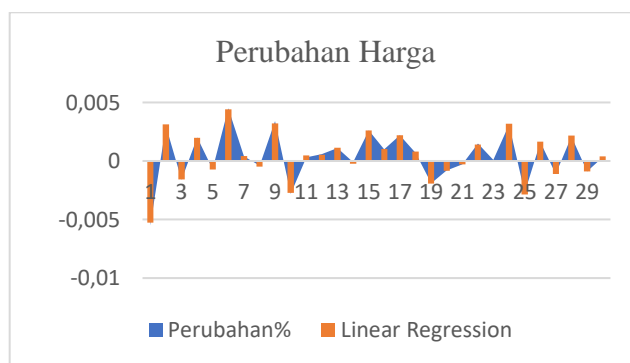
Perubahan%	SVM	kNN	Linear Regression	Fold	Tanggal
0.0036	-0.00125	-0.0005	0.0019532	4	2023-06-06 00:0...
-0.0117	-0.00125	-0.0005	-0.00261362	4	2023-06-07 00:0...
0.0103	-0.0034	-0.00364	0.00526093	2	2023-06-08 00:0...
-0.0007	-0.00125	-0.0005	0.000448448	4	2023-06-09 00:0...
-0.0038	-0.00125	0.00114	-0.00284208	5	2023-06-12 00:0...
-0.0056	-0.00125	0.00226	-0.0112332	1	2023-06-13 00:0...
0.0053	-0.00125	-0.00182	0.00204612	4	2023-06-14 00:0...
-0.0036	-0.00125	0.00054	-0.00166493	1	2023-06-15 00:0...
0.0002	-0.0007	-0.00058	0.00353659	3	2023-06-16 00:0...
0.0037	-0.00125	-0.00406	0.000941299	5	2023-06-18 00:0...
-0.0027	-0.0034	-0.00424	0.000284569	2	2023-06-19 00:0...
-0.0128	-0.0007	-0.00164	-0.0192816	3	2023-06-20 00:0...
-0.0014	-0.00125	-0.00392	-0.00184028	1	2023-06-21 00:0...
-0.0109	-0.00125	-0.00234	-0.00778431	5	2023-06-22 00:0...
0.0031	-0.0034	-0.00564	-0.00095519	2	2023-06-23 00:0...
0.0021	-0.00125	-0.0023	0.00464475	1	2023-06-26 00:0...
-0.0052	-0.00125	0.00254	0.00187183	4	2023-06-27 00:0...
-0.0009	-0.00125	0.00108	0.00106364	5	2023-06-28 00:0...
0.0024	-0.0034	0.00042	-0.00224921	2	2023-06-29 00:0...
0.0060	-0.0007	-0.00102	0.00486178	3	2023-06-30 00:0...
0.0001	-0.0007	-0.00252	-0.00031114	3	2023-07-03 00:0...
0.0023	-0.0007	-0.00252	0.00674232	3	2023-07-04 00:0...
-0.0035	-0.0034	0.00042	-0.00304919	2	2023-07-05 00:0...
-0.0054	-0.00125	0.00146	-0.0108439	1	2023-07-06 00:0...

Gambar 6. Prediksi Menggunakan Regresi Linier.

Pada gambar 6 diatas menunjukkan sampling nilai harga Emas serta hasil prediksi metode Regresi Linier terhadap perubahan nilai harga emas. Data yang ditampilkan selisih perubahan nilai harga, Emas tidak terlalu jauh antara perubahan harga Emas sebenarnya dan perubahan harga Emas berdasarkan prediksi yang dibuat menggunakan metode Regresi Linier.

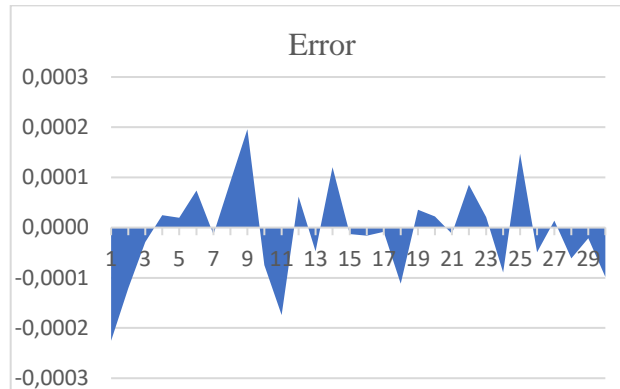


Gambar 7. Grafik Perubahan Harga Hasil Prediksi Regresi Linier. Sumber: Peneliti (2023)



Gambar 8. Grafik Hasil Prediksi Perubahan Harga. Sumber: Peneliti (2023)

Grafik pada Gambar 8 menunjukkan bahwa terdapat selisih nilai/error antara perubahan harga barang yang sesungguhnya dan perubahan harga hasil prediksi dengan menggunakan metode Regresi Linier mendekati nilai sesungguhnya, Namun, ada beberapa harga yang selisihnya jauh dari perubahan harga Regresi Linier sesungguhnya. Hanya 30 titik data yang ditampilkan pada grafik di atas sebagai contoh.



Gambar 9. Grafik Tingkat Error keseluruhan Dataset Sumber: Peneliti (2023)

Grafik error pada Gambar 9 menunjukkan perbedaan antara nilai tukar Emas yang sebenarnya dengan nilai tukar yang diprediksi dengan metode regresi Linear Regresi. Dari data tingkat error tersebut, ditemukan rata-rata tingkat error sekitar 0,000 untuk perubahan harga Emas yang ditentukan dengan menggunakan metode Regresi Linier.

4. KESIMPULAN

Hasil implementasi orange data mining untuk memprediksi harga Emas dapat dilihat di widget Test and Score, yang menggabungkan hasil RMSE dan MAE dari masing-masing model. Nilai RMSE dan MAE dari metode K-Nearest Neighbor (K-NN) sebesar 0.007 dan 0.006, Sedangkan metode Support Vector Machine, memiliki nilai RMSE dan MAE sebesar 0,006 dan 0,005. Berdasarkan nilai RMSE dan MAE tersebut di atas, dan Nilai RMSE dan MAE sebesar 0.004 dan 0.003 dapat disimpulkan bahwa metode Regresi Linier (Linier Regresion) merupakan metode terbaik untuk memprediksi perubahan harga Emas dibandingkan dengan metode K-Nearest Neighbor (K-NN) dan SVM. Metodologi terbaik yang digunakan dalam penelitian ini dapat diterapkan kembali pada penelitian yang akan mendatang. Pada penelitian selanjutnya, disarankan dapat membandingkan metode Regresi Linier dengan metode lainnya untuk memprediksi harga Emas menggunakan tool orange atau dapat menerapkan tools lainnya.

REFERENCES

- [1] M. Owen, V. Vincent, R. Br Ambarita, and E. Indra, "Implementasi Metode Long Short Term Memory Untuk Memprediksi Pergerakan Nilai Harga Emas," *J. Tek. Inf. dan Komput.*, vol. 5, no. 1, p. 96, 2022, doi: 10.37600/tekinkom.v5i1.507.
- [2] R. Kristia Akmal, "Tinjauan Sistematis Untuk Merekomendasi Prediksi Harga Emas," *J. Inov. Inform.*, vol. 7, no. 1, pp. 18–24, 2022, [Online]. Available: <https://jurnal.pradita.ac.id/index.php/jii/article/view/253>
- [3] I. Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "Implementasi Orange Data Mining Untuk Prediksi Harga Bitcoin," *J. Responsif Ris. Sains dan Inform.*, vol. 4, no. 2, pp. 118–125, 2022, doi: 10.51977/jti.v4i2.762.
- [4] A. A. Suryanto, "Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi," *Saintekbu*, vol. 11, no. 1, pp. 78–83, 2019, doi: 10.32764/saintekbu.v11i1.298.
- [5] H. Hozairi, A. Anwari, and S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes," *Netw. Eng. Res. Oper.*, vol. 6, no. 2, p. 133, 2021, doi: 10.21107/nero.v6i2.237.
- [6] F. Ristiano, N. Nurmalarari, and A. Yoraeni, "Impementasi Metode Naive Bayes Untuk Prediksi Harga Emas," *Comput. Sci.*, vol. 1, no. 1, pp. 62–71, 2021, doi: 10.31294/coscience.v1i1.201.
- [7] M. Muharrom, "Jurnal Informatika Dan Teknologi Komputer Komparasi Algoritma Klasifikasi Naive Bayes Dan K-Nearest Neighbors Dalam Analisis Sentimen Terhadap Opini Film Pada Twitter," *Maret*, vol. 3, no. 1, pp. 43–50, 2023.
- [8] M. F. Arfa, M. R. AlFathan, H. B. Lumbantobing, and R. Rahmadenni, "Prediksi Harga Cryptocurrency Dengan Metode Linier Regresi," *SENTIMAS Semin. Nas. Penelit. dan Pengabd. Masy.*, vol. 1, no. 1, pp. 8–15, 2023, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas/article/view/609%0Ahttps://journal.irpi.or.id/index.php/sentimas/article/download/609/332>
- [9] D. Sanajaya, "Analisa Kombinasi Metode Regresi Linier Sederhana dan Single Moving Average pada Harga Beli Emas Digital," vol. XII, no. 1, pp. 196–205, 2022.
- [10] U. Umamah, "Analisis Faktor-Faktor Yang Mempengaruhi Indeks Harga Saham Gabungan Dengan Metode Moderated Regression Analysis," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 8, no. 4, pp. 979–989, 2019, doi: 10.26418/bbimst.v8i4.36772.
- [11] J. Supriyanto, P. Korespondensi, D. Alita, and A. Rahman Isnain, "Penerapan Algoritma K-Nearest Neighbor (K-NN) Untuk Analisis Sentimen Publik Terhadap Pembelajaran Daring," *J. Inform. Dan Rekayasa Perangkat Lunak*, vol. 4, pp. 74–80,

- 2023, [Online]. Available: <https://doi.org/10.33365/jatika.v4i1.2468>
- [12] K. Kartarina, N. K. Sriwinarti, and N. luh P. Juniarti, "Analisis Metode K-Nearest Neighbors (K-NN) Dan Naive Bayes Dalam Memprediksi Kelulusan Mahasiswa," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 3, no. 2, pp. 107–113, 2021, doi: 10.35746/jtim.v3i2.159.
- [13] N. A. Arifuddin, U. Pembangunan, N. Veteran, and S. V. Machine, "Komparasi Naive Bayes dan Support Vector Machine dalam Klasifikasi Jenis Citrus," vol. 22, no. 2, pp. 409–417, 2023.
- [14] S. Y. Pangestu, Y. Astuti, and L. D. Farida, "ALGORITMA SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI SIKAP POLITIK TERHADAP PARTAI Politik Indonesia," *J. Mantik Penusa*, vol. 3, no. 1, pp. 236–241, 2019, [Online]. Available: <https://t.co/eF>
- [15] A. M. Adrian, "Prediksi Menggunakan Algoritma Regresi Linear," *Data Min. Prediksi*, vol. 28, pp. 1–21, 2020.
- [16] N. K. Arkarina, A. W. Widodo, and M. T. Furqon, "Implementasi Regresi Linier Berganda Untuk Prediksi Jumlah Peminat Mata Kuliah Pilihan," *J. Pengemb. Teknol. Inf. Dan Ilmu Komun.*, vol. 3, no. 11, pp. 10462–10467, 2019.
- [17] M. Metode *et al.*, "Prediksi Harga Saham Jakarta Islamic Index," vol. 9, no. 1, pp. 129–135, 2023.
- [18] Reza Maulana and Devy Kumalasari, "Analisis Dan Perbandingan Algoritma Data Mining Dalam Prediksi Harga Saham Ggrm," *J. Inform. Kaputama*, vol. 3, no. 1, pp. 22–28, 2019, [Online]. Available: <https://finance.yahoo.com/quote/GGRM.J>
- [19] A. T. Nurani, A. Setiawan, B. Susanto, D. Salatiga, and J. Tengah, "Perbandingan Kinerja Regresi Decision Tre e dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," vol. 6, no. 1, pp. 34–43, 2023.
- [20] A. R. Wijaya, "Model Prediksi Data Harga Minyak Mentah Dunia Dengan Metode Exponential Smoothing," *Bul. Ilm. Math. Stat. dan Ter.*, vol. 12, no. 1, pp. 21–28, 2023.