Prediksi Harga Saham Bank Rakyat Indonesia Menggunakan Algoritma *Linear Regression* dan *Support Vector Regression*

1st Rizki Nur Annisa
Universitas Buana Perjuangan
Karawang, Indonesia
if19.rizkiannisa@mhs.ubpkarawang.ac.id

2rd Sutan Faisal
Universitas Buana Perjuangan
Karawang, Indonesia
sutan.faisal@ubpkarawang.ac.id

3rd Yana Cahyana
Universitas Buana Perjuangan
Karawang, Indonesia
yana.cahyana@ubpkarawang.ac.id

***Abstract*— Investasi saham merupakan salah satu investasi jangka panjang yang dapat dilakukan oleh masyarakat untuk melindungi risiko biaya terutama untuk mempersiapkan keuangan masa depan. Berinvestasi dalam saham dapat memberikan dividen yang cepat dan cukup besar bagi investor, selain memberikan keuntungan, terdapat juga beberapa faktor yang dapat memengaruhi naik atau turunnya harga saham, seperti kondisi dan kinerja perusahaan, risiko, dividen, suku bunga, kondisi ekonomi, peraturan pemerintah, dan tingkat penurunan. Untuk meminimalkan risiko kerugian bagi investor saat mengambil keputusan investasi, diperlukan analisis secara fundamental dan teknikal. Namun, beberapa orang awam tentang cara menganalisis lebih mendalam tentang perusahaan, maka dari itu cara yang mudah dengan melakukan prediksi dan analisis pergerakan saham. Pada penelitian ini akan dilakukan prediksi harga penutupan saham BRI menggunakan algoritma *Linear Regression* dan *Support Vector Regression* (SVR). Tujuan dari penelitian ini untuk memilih algoritma yang cocok untuk memprediksi harga saham sebagai rekomendasi. Hasil penelitian ini diketahui bahwa algoritma *Linear Regression* dan SVR dapat memprediksi saham, yang menghasilkan nilai *error* RMSE pada algoritma *Linear Regression* sebesar 69.920, sedangkan algoritma SVR menghasilkan nilai *error* RMSE 69.924. Berdasarkan hasil tersebut, diketahui bahwa algoritma *Linear Regression* memiliki performa yang lebih bagus sebagai rekomendasi daripada *Support Vector Regression* dengan selisih nilai *error* sebesar 0.004.**

Kata kunci* — Bank Rakyat Indonesia, Investasi, *Linear Regression*, Prediksi, Saham, *Support Vector Regression

I. PENDAHULUAN

Perusahaan Bank Rakyat Indonesia atau lebih dikenal dengan BRI merupakan sebuah perusahaan swasta dibawah BUMN yang berfokus pada Usaha Mikro Kecil dan Menengah (UMKM) [1]. BRI berfokus pada UMKM karena sebagai sumber dan memiliki potensi pertumbuhan yang sangat tinggi. Perusahaan BRI juga sudah banyak dikenal di masyarakat karena memiliki valuasi yang relative murah dan perusahaan besar dengan *track record* pendapatan yang stabil. Hal tersebut yang membuat saham BRI cenderung stabil dan meningkat.

Investasi saham merupakan salah satu bentuk investasi untuk menyimpan masa kini untuk masa depan. Seiring waktu, berinvestasi dalam saham dapat memberi investor dividen yang cepat dan besar. Dalam hal laba, sejumlah faktor dapat memengaruhi apakah harga saham naik atau turun, seperti kondisi dan kinerja bisnis, risiko, dividen, suku bunga, kondisi ekonomi, peraturan pemerintah, masalah pemerintah atau lainnya, tingkat penurunan, pasokan dan permintaan [2]. Faktor-faktor ini harus dipertimbangkan dan dipelajari oleh investor sebagai strategi investasi dengan melakukan analisis berdasarkan data masa lalu. Analisis ini dapat menjadi informasi yang berguna bagi investor dalam mengambil keputusan. Semakin tinggi pengembalian investor, semakin tinggi risiko yang harus ditanggung investor. Oleh karena itu, untuk meminimalkan risiko kerugian bagi investor saat memutuskan untuk berinvestasi, diperlukan analisis fundamental dan teknikal. Namun, sebagian orang belum mengetahui cara menganalisis perusahaan lebih jauh, maka dari itu cara mudah memprediksi dan menganalisis pergerakan saham dengan metode yang lebih akurat untuk mulai memahami saham. Estimasi harga saham juga berguna bagi investor untuk melihat prospek investasi ke depan [3].

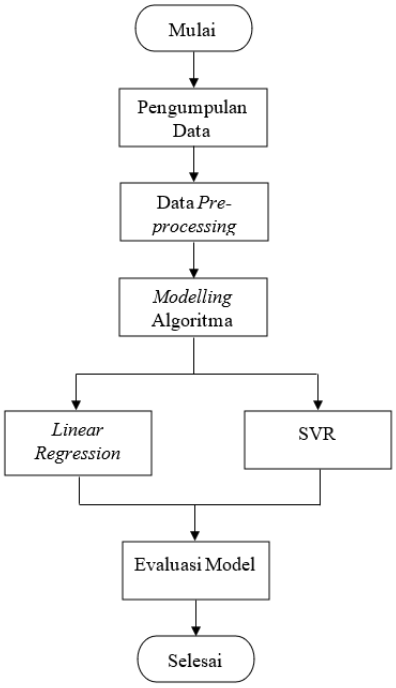
Keren & Oetama [4] melakukan penelitian untuk meramalkan outlook harga saham BCA periode 2015-2017 menggunakan algoritma regresi linier dan SPSS 16.0. Hasil penelitian berupa perhitungan MAPE (*Moving Average Percentage Error*) dengan nilai terendah yang terjadi pada bulan Mei 2017 sebesar 234,64. Bergantung pada hasil pencarian, kesalahan rata-rata yang ditampilkan berfluktuasi. Ramdhani & Mubarok [5] juga melakukan penelitian prediksi harga penutupan saham PT Aneka Tambang Tbk (PT Antm menggunakan SVR (*Support Vector Regression*). Penelitian ini menggunakan 262 data record. Hasil penelitiannya berupa nilai RMSE memiliki nilai terendah sebesar 22.662. Penelitian selanjutnya oleh Novianti et al. [6] terkait prediksi konsumsi bahan bakar PT Indonesia Power dengan menggunakan regresi support vector (SVR). Penelitian ini menggunakan data dari tanggal 21 Agustus 2020 sampai dengan 31 Agustus 2021. Hasil penelitian menggunakan kernel polynomial menunjukkan dapat memprediksi dengan baik dengan nilai RMSE sebesar 0,0266 dan memiliki nilai MAPE kurang dari 10%.

Berdasarkan penelitian sebelumnya, pemodelan prediktif dengan menggunakan metode regresi memiliki pengaruh yang besar terhadap pengambilan keputusan. Algoritma regresi linier memberikan akurasi yang tinggi dalam kasus prediksi, algoritma SVR juga banyak digunakan untuk prediksi presisi tinggi dan akan digunakan sebagai algoritma pembanding. Oleh karena itu, penelitian ini bertujuan untuk menentukan model prediksi harga saham terbaik dengan menerapkan regresi linier dan algoritma SVR, dengan judul "Prediksi Harga Saham Bank Rakyat Indonesia Menggunakan Algoritma *Linear Regression* dan *Support Vector Regression*".

II. METODOLOGI PENELITIAN

A. Alur Penelitian

Penelitian ini menggunakan data harga saham BRI untuk diprediksi dengan dua metode yaitu algoritma *Linear Regression* dan SVR. Dataset akan diolah dibagian *pre-processing* sebelum diprediksi. Dataset yang sudah melewati tahap pembersihan, dilakukan modeling dengan algoritma *Linear Regression* dan SVR, kemudian dilakukan evaluasi terhadap performa algoritma masing-masing. Alur penelitian dapat dilihat pada Gambar 1 berikut.



Gambar 1 Alur Penelitian

B. Sumber Data

Dataset yang digunakan pada penelitian ini berupa data saham Bank Rakyat Indonesia. Data saham yang digunakan diakses dan diunduh dari *Website Yahoo Finance* pada 02 Januari 2023. Dataset yang digunakan selama rentang 5 tahun dari tahun 02 Januari 2018 hingga 30 Desember 2022. Dataset terdiri dari 1.235 total data dengan tujuh variabel yaitu *date*, *open*, *high*, *low*, *close*, *adj close*, dan *volume*.

C. *Linear Regression*

Linear Regression merupakan salah satu jenis analisis prediktif. Model regresi sering digunakan dengan cara yang sama untuk keputusan praktis dan penelitian teoretis [7]. Regresi linier dalam pemodelan dilakukan berdasarkan data sebelumnya. Regresi bertujuan untuk meminimalkan selisih antara nilai prediksi dengan nilai sebenarnya, dengan memodelkan data sehinggga regresi dapat dipandang sebagai alat ukur korelasi antar variabel [8]. Data regresi linier terdiri dari dua variabel yaitu x (variabel bebas) dan y (variabel terikat). Model kemudian dibangun untuk memprediksi nilai y pada nilai x tertentu. Model persamaan regresi linear [9], didefinisikan pada persamaan (1), (2), dan (3) berikut.

$$y = a + bx \quad (1)$$

$$a = \frac{\sum y.\sum x^2 - \sum x.\sum xy}{n\sum x^2 - (\sum x)^2} \quad (2)$$

$$b = \frac{n\sum xy - \sum x.\sum y}{n\sum x^2 - (\sum x)^2} \quad (3)$$

Dalam persamaan (1), (2), dan (3), “y” mewakili variabel terikat, “x” sebagai variabel bebas, “a” dan “b” adalah nilai intercept dan koefisien dari model, dan “n” adalah total data.

D. *Support Vector Regression*

Support Vector Regression adalah algoritma pengembangan SVM (*Support Vector Machine*). Algoritma SVR diterapkan pada permasalahan regresi yang memberikan nilai dalam bentuk kontinu atau real [10]. SVR bertujuan untuk menemukan batas atau jalur super terbaik. Hal ini menjadikan SVR sebagai algoritma yang menghasilkan nilai prediksi yang mendekati sempurna dengan tingkat kesalahan yang kecil dan akurasi yang baik.

Dalam beberapa metode analisis *data mining*, banyak yang menggunakan fungsi linear, namun ada juga kasus non linier. Kasus non-linier dapat diatasi dengan mengubah data ke dimensi yang lebih tinggi. Metode ini menggunakan metode inti, dimana sebuah data x pada input *space* didekomposisi secara linear dan dipetakan ke dalam *feature space* baru [11]. Menurut Muhammad Romzi & Kurniawan [12], fungsi perkalian yang digunakan dalam metode SVR dapat dilihat pada Tabel 1.

Tabel 1 Fungsi Kernel SVR

Jenis Kernel	Definisi Rumus
Linear	$K(x_i, x) = x_i, x$
Radial Basis Function (RBF)	$K(x_i, x) = exp\,(\frac{\|\|x_i, x\|\|^2}{-2\sigma^2})\gamma = \frac{1}{2\sigma^2}$
Polynomial	$K(x_i, x) = (x_i^T x_j + t)^d$

E. *Root Mean Square Error* (RMSE)

Dilakukan evaluasi terhadap model algoritma untuk mengukur keakuratan algoritma. Evaluasi diperoleh dari perhitungan *root mean square error* (RMSE). *Root mean square error* (RMSE) adalah perkiraan dengan jumlah kesalahan kuadrat atau perbedaan antara nilai aktual dan nilai prediksi. Semakin kecil nilai RMSE, semakin akurat algoritma yang digunakan. Berdasarkan hasil evaluasi, akan didapatkan model algoritma terbaik untuk penelitian ini. Rumus RMSE didefinisikan pada persamaan (4) sebagai berikut.

$$RMSE = \sqrt{\sum \frac{(Y' - Y)^2}{n}} \quad (4)$$

III. HASIL DAN PEMBAHASAN

A. *Data Pre-Processing*

1.

Pre-processing data dilakukan dengan tujuan mengubah data mentah menjadi data yang lebih bersih dan efisien. Tahapan-tahapan yang dilakukan pada data *pre-processing* berupa menghapus atribut yang tidak digunakan, menghapus nilai *missing value* pada data, dan merubah tipe data yang digunakan.

1. Reduksi Data

Pada tahapan ini dilakukan pengurangan jumlah data saham Bank Rakyat Indonesia dengan memilih atribut yang relevan untuk digunakan pada penelitian ini, atribut yang tidak relevan atau tidak digunakan akan dihapus. Pada penelitian ini, data yang digunakan sebesar 1.235 data dengan 7 atribut yaitu *date*, *open*, *high*, *low*, *close*, *adj. close*, dan *volume*. Setelah melakukan tahap reduksi, atribut yang dibutuhkan dalam penelitian ini terdapat 3 atribut yaitu *date*, *open*, dan *close*. Hasil data sebelum dan sesudah direduksi dengan kode program *python* seperti Tabel 2 berikut.

Tabel 2 Proses Reduksi Data

Sebelum	Sesudah
Atribut	**Atribut**
Date, Open, High, Low, Close, Adj. Close, Volume	*Date, Open, Close*

2. Pembersihan Data

Pada proses ini untuk menghilangkan data yang tidak revelan, nilai kosong, dan untuk memperbaiki kesalahan pada data dengan kode program *python*. Pada dataset penelitian ini yang diambil dari situs *yahoo finance*, yang terdiri dari 1.235 data terdapat 1 kesalahan data atau nilai *null* yang dapat mempengaruhi proses prediksi saham sehingga data tersebut akan diganti nilainya dengan nilai mean dari seluruh data yang ditunjukkan pada Tabel 3.

Tabel 3 Proses Pembersihan Data

Sebelum		Sesudah	
Atribut	***Missing Value***	**Atribut**	***Missing Value***
Date	2019-06-19	*Date*	2019-06-19
Open	*Null*	*Open*	3941
Close	*Null*	*Close*	3936

3. Transformasi Data

Proses transformasi data dengan kode program *python* bertujuan untuk mengubah bentuk suatu data ke data yang dibutuhkan untuk proses prediksi saham. Tranformasi data dilakukan untuk merubah format data *date* dari format "hari/bulan/tahun" menjadi format "tahun-bulan-hari" dan transformasi jenis tipe data atribut *open* dan *close* dari float menjadi integer untuk memudahkan proses prediksi seperti Tabel 4.

Tabel 4 Proses Transformasi Data

Sebelum		Sesudah	
Atribut	***Missing Value***	**Atribut**	***Missing Value***
Date	02/01/2018	*Date*	2018-01-02
Open	*float64*	*Open*	*int64*
Close	*float64*	*Close*	*int64*

B. Pemodelan Algoritma

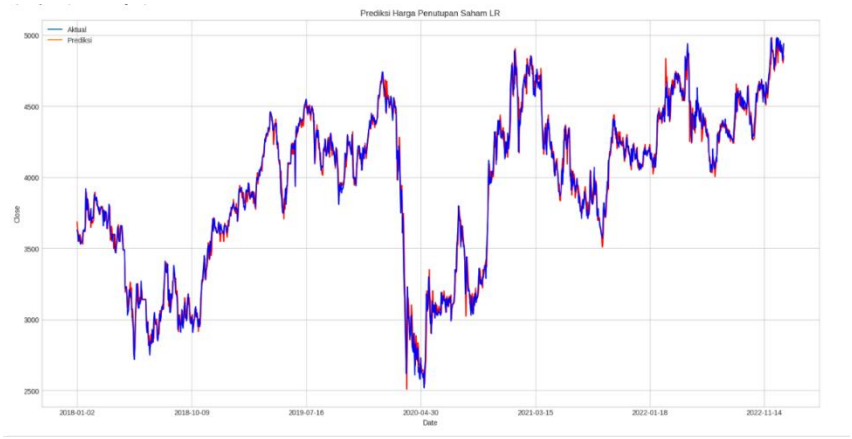
Pada tahapan ini, pemodelan dilakukan dengan dua algoritma yaitu *Linear Regression* dan SVR menggunakan pemrograman *python*. Hasil pemodelan berupa visualisasi dan prediksi harga penutupan saham.

1. Pemodelan *Linear Regression*

Prediksi harga penutupan saham BRI menggunakan pemrograman *python* dengan algoritma *linear regression* menghasilkan model persamaan algortima yang didapatkan dari nilai koefisien dan intercept yang ditunjukkan pada persamaan (5).

$$y = 36.53831967 + 0.98943586x \quad (5)$$

Dari hasil model persamaan didapatkan hasil prediksi yang divisualisasikan dengan grafik yang ditunjukkan pada Gambar 2.



Gambar 2 Visualisasi Prediksi LR

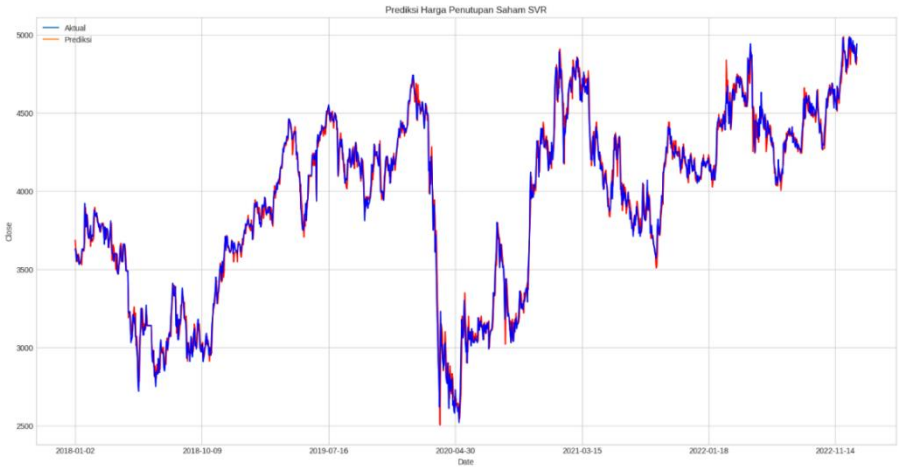
Pada grafik hasil visualisasi Gambar 2 dengan algoritma *linear regression*, terlihat bahwa dari tahun 2018-2022 harga saham mengalami fluktuasi. Pola yang diambil dari hasil prediksi dari tahun 2018 hingga 2019 mengalami kenaikan yang signifikan sebesar 23.27%, kemudian tahun 2019 ke tahun 2020 mengalami penurunan prediksi saham sebesar -13.37%, kemudian tahun 2020-2021 mengalami kenaikan sebesar 17.95%, dan 2021-2022 mengalami kenaikan 6.28% atau bisa dikatakan mengalami penurunan dari tahun sebelumnya, jadi untuk tahun selanjutnya yaitu 2023 diprediksi mengalami kenaikan dengan mengesampingkan faktor luar yang lain dengan rata-rata kenaikan sebesar 8.53%. Hasil analisis tersebut dapat menjadi referensi harga penutupan saham untuk masa yang akan mendatang. Berikut persentase prediksi kenaikan dan penurunan harga penutupan saham yang ditunjukkan pada Tabel 5.

Tabel 5 Persentase Prediksi LR

Tahun	Persentase	Keterangan
2018-2019	23.27%	Naik
2019-2020	-13.37%	Turun
2020-2021	17.95%	Naik
2021-2022	6.28%	Turun
Rata-rata	**8.53%**	**Naik**

2. Pemodelan *Support Vector Regression*

Prediksi harga penutupan saham BRI menggunakan pemrograman *python* dengan algoritma SVR menggunakan kernel 'linear' dengan parameter C 1.0 dan epsilon 0.1 menghasilkan pola prediksi yang divisualisasikan menggunakan grafik seperti pada Gambar 3.



Gambar 3 Visualisasi Prediksi SVR

Pada grafik hasil visualisasi Gambar 3 dengan algoritma SVR, terlihat bahwa dari tahun 2018-2022 harga saham mengalami fluktuasi. Pola yang diambil dari hasil prediksi menggunakan kernel linear dihitung per-tahun. Untuk per-tahunnya, dari tahun 2018 hingga 2019 mengalami kenaikan yang signifikan sebesar 23.36%, kemudian tahun 2019 ke tahun 2020 mengalami penurunan prediksi saham sebesar -13.41%, kemudian tahun 2020-2021 mengalami kenaikan sebesar 18.01%, dan 2021-2022 mengalami kenaikan 6.30% atau bisa dikatakan mengalami penurunan dari tahun sebelumnya, jadi untuk tahun selanjutnya yaitu 2023 diprediksi mengalami kenaikan dengan mengesampingkan faktor luar yang lain dengan rata-rata kenaikan sebesar 8.56%. Hasil analisis tersebut dapat menjadi referensi harga penutupan saham untuk masa yang akan mendatang. Berikut persentase prediksi kenaikan dan penurunan harga penutupan saham yang ditunjukkan pada Tabel 6.

Tabel 6 Persentase Prediksi SVR

Tahun	Persentase	Keterangan
2018-2019	23.36%	Naik
2019-2020	-13.41%	Turun
2020-2021	18.01%	Naik
2021-2022	6.30%	Turun
Rata-rata	**8.56%**	**Naik**

C. Perbandingan Algoritma

Dari hasil pemodelan kedua algoritma pada data saham Bank Rakyat Indonesia menggunakan pemrograman *python* didapatkan hasil evaluasi menggunakan RMSE yang dapat dilihat pada Tabel 7

Tabel 7 Perbandingan Performa

Algoritma	RMSE
Linear Regression	**69.920**
Support Vector Regression	69.924

Hasil penelitian menunjukkan bahwa kedua algoritma mampu memprediksi data saham Bank Rakyat Indonesia dengan baik dengan nilai RMSE pada evaluasi model, hasil RMSE digunakan untuk menilai seberapa besar selisih dari data aktual dengan data prediksi dari perhitungan akar kuadrat rata-rata kesalahan *error*. Semakin mendekati angka 0 nilainya, maka selisihnya kecil. RMSE dari kedua algoritma sebesar diangka 69 atau selisih prediksi dengan data aktual sejauh 2 digit dengan menghasilkan nilai RMSE algoritma *Linear Regression* sebesar 69.920 lebih baik dalam memprediksi daripada SVR dengan nilai 69.924, karena memiliki nilai RMSE lebih kecil dengan selisih sebesar 0.004. Hasil evaluasi dari model kedua algoritma menghasilkan selisih aktual dan prediksi yang cukup besar.

IV. PENUTUP

A. Kesimpulan

Hasil penelitian menunjukkan bahwa kedua algoritma dapat memprediksi harga penutupan saham BRI selama periode 2018-2019 dengan rata-rata pola kenaikan model *Linear Regression* sebesar 8.53% dan SVR sebesar 8.56%. Evaluasi model juga menunjukkan bahwa algoritma *Linear Regression* lebih direkomendasikan pada penelitian prediksi saham karena memiliki performa yang lebih baik dengan nilai RMSE 69.920 lebih kecil 0.004 daripada SVR sebesar 69.924.

B. Saran

Pada penelitian ini berfokus pada model prediksi penutupan saham. Pada penelitian yang akan datang diharapkan dapat mengembangkan model prediksi menjadi sebuah sistem atau aplikasi yang dapat bermanfaat bagi masyarakat untuk melihat prediksi pergerakan harga saham

2. PENGAKUAN

Naskah Ilmiah ini adalah sebagian dari penelitian Tugas Akhir milik Rizki Nur Annisa dengan judul Implementasi *Linear Regression* dan *Support Vector Regression* dalam Memprediksi Harga Saham Bank Rakyat Indonesia yang dibimbing oleh Bapak Sutan Faisal dan Bapak Yana Cahyana.

3. DAFTAR PUSTAKA

[1] Investor Daily, "4 Alasan Mengapa Investor Harus Beli Saham BBRI," *Gora Kunjana*, 2020. https://investor.id/market-and-corporate/221262/4-alasan-mengapa-investor-harus-beli-saham-bbri

[2] J. S. Putra, R. D. Ramadhani, and A. Burhanuddin, "Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regresion Sebagai Strategi Jual Beli Saham," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 1, pp. 1–10, 2022, doi: 10.20895/dinda.v2i1.273.

[3] R. Maulana and D. 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

[4] M. Keren and R. S. Oetama, "Prediksi prospek harga saham perusahaan perbankan Menggunakan Regresi Linear (Study Kasus Bank BCA Tahun 2015-2017)," *JSI J. Sist. Inf.*, vol. 11, no. 1, pp. 1703–1710, 2019, doi: 10.36706/jsi.v11i1.8105.

[5] Y. Ramdhani and A. Mubarok, "Analisis Time Series Prediksi Penutupan Harga Saham," *J. Responsif*, vol. 1, no. 1, pp. 77–82, 2019, doi: https://doi.org/10.51977/jti.v1i1.92.

[6] F. Novianti, N. Ulinnuha, M. Hafiyusholeh, and A. Arianto, "Prediksi Penggunaan Bahan Bakar pada PLTGU menggunakan Metode Support Vector Regression (SVR)," *Techno.Com*, vol. 21, no. 2, pp. 249–255, 2022, doi: 10.33633/tc.v21i2.5712.

[7] T. M. H. Hope, "Linear regression," *Mach. Learn. Methods Appl. to Brain Disord.*, pp. 67–81, 2019, doi: 10.1016/B978-0-12-815739-8.00004-3.

[8] A. Hidayanti, A. M. Siregar, S. A. P. Lestari, and Y. C. Cahyana, "Model Analisis Kasus Covid-19 Di Indonesia Menggunakan Algoritma Regresi Linier Dan Random Forest," *Petir*, vol. 15, no. 1, pp. 91–101, 2021, doi: 10.33322/petir.v15i1.1487.

[9] N. R. Setyoningrum and P. J. Rahimma, "Implementasi Algoritma Regresi Linear Dalam Sistem Prediksi Pendaftar Mahasiswa Baru Sekolah Tinggi Teknologi Indonesia Tanjungpinang," *Pros. Semin. Nas. Ilmu Sos. dan Teknol.*, no. 4, pp. 13–18, 2022.

[10] M. P. Lestari *et al.*, "Peramalan Pertambahan Pasien Covid-19 Menggunakan Support Vector Regression Forecasting Growth of Covid-19 Patients Using Support," vol. 8, no. 5, pp. 9497–9507, 2021.

[11] A. Prahutama, T. W. Utami, and H. Yasin, "Prediksi Harga Saham Menggunakan Support Vector Regression Dengan Algoritma Grid Search," *Media Stat.*, vol. 7, no. 1, pp. 29–35, 2014, doi: https://doi.org/10.14710/medstat.7.1.29-35.

[12] Muhammad Romzi and B. Kurniawan, "Pembelajaran Pemrograman *Python* Dengan Pendekatan Logika Algoritma," *JTIM J. Tek. Inform. Mahakarya*, vol. 03, no. 2, pp. 37–44, 2020.

ISSN: 2715-2766
Vol. 6 No: 1, Januari 2025