

## Prediksi Harga Saham Indonesia pada Masa Covid-19 Menggunakan Regresi Pohon Keputusan

Rina Nopianti<sup>1\*</sup>, Andreas Tri Panudju<sup>2</sup>, Angrian Permana<sup>3</sup>

<sup>1</sup> Universitas Bina Bangsa, rinanopianti.binabangsa@gmail.com

<sup>2</sup> Universitas Bina Bangsa, Panudju2002@gmail.com

<sup>3</sup> Universitas Bina Bangsa, mr.angrianpermana@gmail.com

### ABSTRACT

*Stock price prediction is an interesting area of data mining. Many variables affect stock prices. Stock prices fluctuate, especially in the Covid-19 era which has an impact on the economy. This study aims to predict the stock price of Indonesian telecommunications, especially in the covid19 era. The object of this research is the financial statements of companies listed on the Indonesia Stock Exchange, where data related to stock prices of telecommunications companies are used as samples in the data processing of this study. In this study, the regression techniques that will be used are Multiple Linear Regression, Support Vector Regression, Decision Tree Regression, and K-Nearest Regression. The results showed that the fundamental data and stock prices have a relationship. High correlation coefficient resulted from Decision Tree Regression and K-Nearest Regression. Decision Tree Regression produces good results on the Train Test Split and KFold Cross Validation data, 2.99% and 2.98% repeatedly. It can conclude that despite the fact that the pandemic scenario has had a significant impact on the stock market, the study's findings suggest that fundamental data and stock prices have a relationship. Decision Tree Regression and K-Nearest Regression both produced high correlation coefficients.*

**Keywords:** Stock Price, Prediction, Decision Tree, Regression, COVID-19

### ABSTRAK

Prediksi harga saham adalah area Data Mining yang menarik. Banyak variabel yang mempengaruhi harga saham. Harga saham menjadi fluktuatif, terutama di era covid-19 yang berdampak pada perekonomian. Penelitian ini bertujuan untuk memprediksi harga saham telekomunikasi Indonesia khususnya pada era covid19. Obyek penelitian ini adalah laporan keuangan perusahaan yang terdaftar di Bursa Efek Indonesia, dimana data-data terkait harga saham dari perusahaan telekomunikasi dijadikan sampel dalam pengolahan data penelitian ini. Dalam penelitian ini teknik regresi yang akan digunakan adalah Regresi Linier Berganda, *Regresi Support Vector*, Regresi Pohon Keputusan, dan Regresi *K-Nearest*. Hasil penelitian menunjukkan bahwa data fundamental dan harga saham memiliki hubungan. Koefisien korelasi yang tinggi dihasilkan dari Regresi Pohon Keputusan dan *Regresi K-Nearest*. Regresi Pohon Keputusan menghasilkan hasil yang baik pada data *Train Test Split* dan *KFold Cross Validation*, 2,99% dan 2,98% secara berulang. Dapat disimpulkan bahwa terlepas dari kenyataan bahwa skenario pandemi memiliki dampak yang signifikan terhadap pasar saham, temuan studi menunjukkan bahwa data fundamental dan harga saham memiliki hubungan. Regresi Pohon Keputusan dan Regresi *K-Nearest* keduanya menghasilkan koefisien korelasi yang tinggi.

**Kata Kunci :** Harga Saham, Prediksi, Pohon Keputusan, Regresi, COVID-19

*Naskah diterima: 13-10-2021, direvisi: 10-01-2022, diterbitkan: 01-04/2022*

### PENDAHULUAN

Saat ini masyarakat Indonesia sudah concern terhadap investasi. Investasi umum adalah membeli saham, uang, tanah, dan emas. Investasi adalah kegiatan untuk mendapatkan keuntungan, sedangkan saham merupakan bukti kepemilikan perusahaan yang menarik untuk diinvestasikan. Ada

banyak faktor yang mempengaruhi harga saham. Faktor rasional dan berbagai faktor irasional menjadi faktor penentu dalam pembelian saham (Purnama, 2017). Saat ini, artikel berita dan media sosial juga dapat mempengaruhi keputusan investor.

*Coronavirus Disease 2019 (COVID-19)* menyebar dengan mudah dan mematikan

bagi sebagian orang, dan penyebabnya masih terus ditemukan. Pada 12 Maret 2020, (Organisasi Kesehatan Dunia) WHO mengumumkan wabah tersebut sebagai pandemi. Pandemi menyebar dengan cepat, Cina, Korea Selatan, Italia, Inggris, Amerika Serikat, dan negara-negara lain telah terinfeksi. Mereka memiliki kasus positif yang sangat besar sebelum mereka dapat mengendalikan wabah. Mengontrol wabah berarti mengendalikan pergerakan orang, pada awalnya. Artinya, gaya hidup, konsumsi, dan belanja masyarakat berubah. Harga saham turun secara global. Pandemi global COVID-19 menyebabkan pola pasar saham yang berbeda sebelum dan sesudah berita tersebut hits (D. Zhang et al., 2020).

Kasus pertama COVID-19 di Indonesia diumumkan pada Maret 2020. Penemuan kasus tersebut mengakibatkan kondisi yang tidak stabil baik dalam perekonomian maupun keamanan masyarakat. Situasi saham di Indonesia pada saat COVID-19 dapat dilihat dari Indeks Harga Saham Gabungan (IHSG) yang merupakan harga saham gabungan Indonesia.

Prediksi harga saham dapat membantu investor dalam mengambil keputusan atas transaksi pasar saham. Prediksi harga saham digunakan sebagai bahan pertimbangan dalam memutuskan apakah akan mempertahankan atau menjual sahamnya. Ada dua pendekatan untuk menganalisis pergerakan harga saham, yaitu analisis fundamental dan teknikal.

Di Indonesia, pergerakan masyarakat juga terbatas akibat COVID-19, namun harus tetap terkoneksi. Salah satu sektor yang masih bertahan dengan kinerjanya adalah telekomunikasi. Penelitian ini akan mengamati saham data perusahaan telekomunikasi yang akan diprediksi menggunakan teknik regresi termasuk era Covid19. Kontribusi dari makalah ini adalah kami bertujuan untuk mendapatkan model terbaik untuk menggeneralisasi inferensi setiap keputusan prediktif saham masing-masing perusahaan dan memiliki *Mean Absolute Percentage Error* (MAPE) terkecil.

## KAJIAN LITERATUR

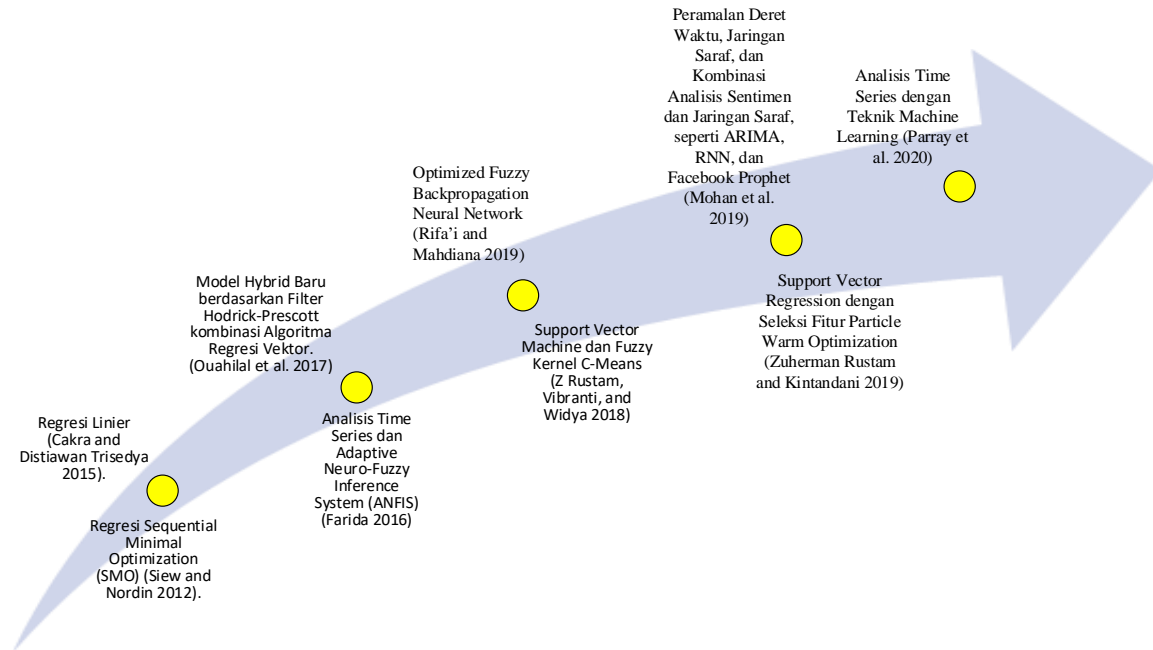
Banyak penelitian telah dilakukan di bidang prediksi harga saham. Anwar dan Deni (Rifa'i & Mahdiana, 2019) memprediksi Indeks Harga Saham Gabungan Bursa Efek Indonesia menggunakan *Optimized Fuzzy Backpropagation Neural Network* dan juga menggunakan Algoritma Genetika. Metode ini berhasil memprediksi nilai MAPE sebesar 8,42% dari indeks harga saham gabungan. Support vector machine dan fuzzy Kernel C-Means untuk memprediksi arah pergerakan harga saham Indonesia, dimana didapatkan hasilnya bagus karena nilai *Normalized Mean Squared Error* (NMSE) yang relatif kecil (Z Rustam et al., 2018).

Model peramalan deret waktu, jaringan saraf, dan kombinasi analisis sentimen dan jaringan saraf, seperti ARIMA, RNN, dan Facebook Prophet untuk memprediksi harga saham. Hasil terbaik yang dicapai adalah dengan RNN (Mohan et al., 2019). Harga saham dan informasi teks dari korelasi berita juga ditemukan. Yuniar (Farida, 2016) juga menggunakan analisis time series dan *Adaptive Neuro-Fuzzy Inference System* (ANFIS) untuk memprediksi harga saham mingguan. Parray et al (Parray et al., 2020) menganalisis pergerakan harga saham menggunakan analisis *time series* dengan teknik *machine learning*. Akurasi meningkat sekitar 2% dengan model yang diusulkan.

Salah satu teknik yang digunakan dalam memprediksi tren saham adalah teknik regresi oleh Siew dan Nordin (Siew & Nordin, 2012). Penelitian ini menghasilkan teknik *Regresi Sequential Minimal Optimization* (SMO) mengungguli teknik regresi lainnya. Yahya dan Bayu (Cakra & Distiawan Trisedya, 2015) menganalisis beberapa harga saham Indonesia dan analisis sentimennya menggunakan Regresi Linier. Penelitian lain yang dilakukan oleh Zuherman dan Puteri (Zuherman Rustam & Kintandani, 2019) memprediksi beberapa harga pasar saham dengan menggunakan *Support Vector Regression* dengan Seleksi Fitur *Particle Warm Optimization*, menghasilkan kinerja yang baik.

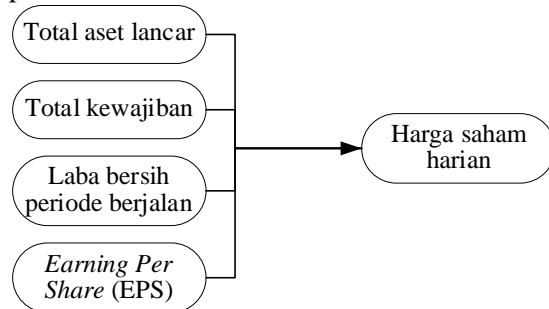
Ouahilal et al (Ouahilal et al., 2017) mengoptimalkan harga pasar saham menggunakan model *hybrid* baru berdasarkan filter *Hodrick-Prescott* dan mendukung algoritma regresi vektor. Hasilnya menunjukkan akurasi prediksi harga pasar saham yang sangat baik dengan

waktu eksekusi yang sangat minimalis. Asghar et al (Asghar et al., 2019) mengembangkan sistem prediksi tren pasar saham menggunakan regresi berganda. Dalam penelitian ini, teknik regresi akan digunakan untuk memprediksi harga pasar saham.



Gambar 1. Perkembangan penelitian pada topik prediksi harga saham

Dalam penelitian ini, kami menggunakan data empat tahun untuk memprediksi harga pasar saham. Sektor telekomunikasi dipilih karena peran untuk menjaga masyarakat tetap terhubung melalui pandemi. Empat perusahaan telekomunikasi di Indonesia akan diamati. Kerangka pikir penelitian ini adalah:

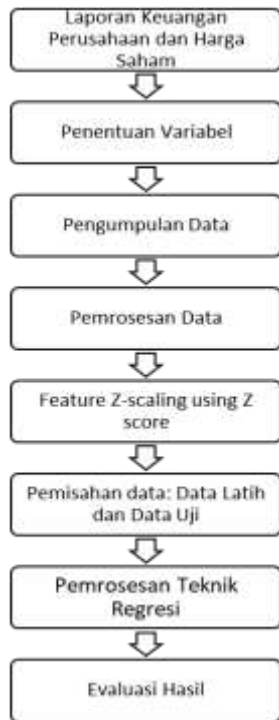


Gambar 2. Kerangka Pikir Penelitian

## METODE PENELITIAN

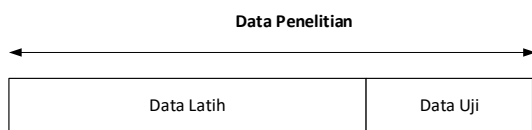
### Metode Penelitian

Metodologi yang digunakan dalam penelitian ini dibagi menjadi delapan langkah seperti yang terlihat pada Gambar 3. Obyek penelitian ini adalah laporan keuangan perusahaan yang terdaftar di Bursa Efek Indonesia, dimana data-data terkait harga saham dari perusahaan telekomunikasi dijadikan sampel dalam pengolahan data penelitian ini. Setelah mengumpulkan laporan keuangan perusahaan dan harga saham, langkah selanjutnya adalah menentukan variabel. Variabel independen adalah total aset lancar, total kewajiban, laba bersih periode berjalan, dan *Earning Per Share* (EPS). Variabel terikatnya adalah harga saham harian penutupan.



Gambar 3. Metodologi Penelitian

Langkah selanjutnya adalah mengumpulkan data. Laporan keuangan dirilis setiap triwulan biasanya pada tanggal 31 Desember, 31 Maret, 30 Juni, dan 30 September. Pengumpulan data perlu dilakukan dengan hati-hati karena dampak laporan keuangan dapat terlihat setelah dirilis. Misalnya, fundamental perusahaan yang akan memandu investor di bulan Januari adalah laporan keuangan Q4 perusahaan dari tahun sebelumnya.

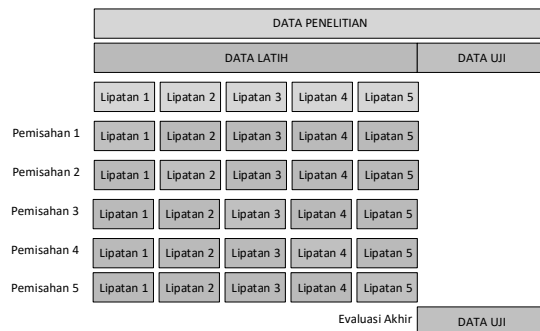


Gambar 4. Pemisahan Data

Setelah data terkumpul, diperlukan pembersihan data dalam preprocessing data. Proses selanjutnya adalah penskalaan fitur dengan *Zscore*. Penskalaan fitur adalah langkah standarisasi data. Ini diterapkan pada variabel independen atau fitur data yang akan dinormalisasi dalam rentang tertentu. *Zscore* atau skor standar membantu menghitung probabilitas skor yang terjadi dalam distribusi normal dan

membandingkan dua skor dari distribusi normal yang berbeda.

Setelah memproses data, data harus dipecah menjadi set data pelatihan dan set data pengujian. Training set akan membuat model menggunakan teknik regresi. Satu set tes akan digunakan untuk mengevaluasi hasil. Pemisahan uji latih biasa dan Validasi Silang *K-Fold* digunakan untuk bereksperimen dengan kumpulan data (G. Zhang et al., 2017). Kedua metode validasi sampling tersebut bertujuan untuk mengetahui kekokohan model yang berhasil dibuat. Perbedaan antara kedua pemisahan



data tersebut terlihat pada Gambar 4 dan Gambar 5.

Gambar 5. Validasi Silang K-Fold (Selvamuthu et al., 2019)

Dalam penelitian ini teknik regresi yang akan digunakan adalah Regresi Linier Berganda, *Regresi Support Vector* (Yu et al., 2014), Regresi Pohon Keputusan, dan Regresi *K-Nearest* (Chen & Hao, 2017). *Multiple Linear Regression* (MLR) terdiri dari lebih dari satu variabel prediktor  $x$  dan variabel respon  $Y$  (Olive, 2017). MLR dirumuskan dengan formula (1).

$$Y_i = x_{i,1}\beta_1 + x_{i,2}\beta_2 + \dots + x_{i,n}\beta_n \dots (1)$$

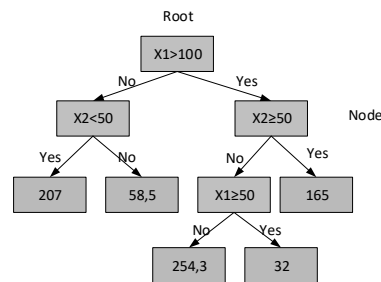
*Support Vector Regression* (SVR) adalah model hyperplane mendukung vektor dengan  $\epsilon$ -tube, daerah  $\epsilon$ -insensitive (Boujelbene et al., 2010). SVR memiliki konsep yang hampir sama dengan *Support Vector Machines* (SVM). SVM dapat

membuat prediksi berdasarkan fungsi linier dalam ruang fitur berdimensi tinggi. Algoritma didasarkan pada teori optimasi dengan menerapkan pembelajaran untuk menemukan asal-usulnya dalam studi teori statistik. Perbedaan antara SVR dan SVM adalah pada tabungnya. Tabung berkontribusi untuk mengoptimalkan pemecahan masalah dengan menemukan yang terbaik mendekati fungsi bernilai kontinu (Chen & Hao, 2017).

Regresi Pohon Keputusan kurang populer dibandingkan dengan pohon klasifikasi. Namun, dibandingkan dengan algoritma lain, sangat kompetitif dan biasanya digunakan untuk memecahkan masalah dalam kehidupan nyata (Czajkowski & Kretowski, 2016). Regresi Pohon Keputusan yang dibangun dari *Root*, *Node*, dan *Leaf* mirip dengan pohon klasifikasi. Perbedaan antara pohon regresi dan pohon klasifikasi adalah pada outputnya. Output pohon regresi adalah nilai kontinu atau nilai bilangan real alih-alih kelas (Nguyen et al., 2016). Untuk beberapa masalah, kelemahan dari pohon keputusan adalah overfitting dan tidak stabil. Keuntungan dari pohon keputusan adalah mudah dipahami, diinterpretasikan, berguna dalam eksplorasi data, dan lebih sedikit persiapan data yang diperlukan (Xianya et al., 2019). Contoh sederhana regresi pohon keputusan dapat dilihat pada Gambar 5.

*K-Nearest Neighbor Regression* (KNN-R) mengungguli algoritma pembelajaran mesin lainnya (Vanneschi et al., 2020). KNN-R bekerja mirip dengan klasifikasi KNN, perbedaannya terletak pada outputnya. Sedangkan output dari klasifikasi KNN adalah kelas yang termasuk ke dalam tetangga terdekat, output KNN-R adalah rata-rata dari nilai k tetangga terdekat. Rumus jarak KNN (Cai et al., 2018) ditunjukkan pada formula (2). Nilai jarak baik 1, jarak Manhattan atau 2, jarak *Euclidean*.

$$d(X, Y) = (\sum_{i=1}^n |x_i - y_i|^p)^{1/p} \dots\dots\dots (2)$$



Gambar 6. Contoh Regresi Pohon Keputusan (Xianya et al., 2019)

Langkah terakhir adalah mengevaluasi hasil. Koefisien korelasi dan *Mean Absolute Percentage Error* (MAPE) akan digunakan untuk mengevaluasi hasil (Agustin et al., 2016). Hasil yang bagus dengan koefisien korelasi yang tinggi dan jumlah MAPE yang minimum diharapkan. Rumus MAPE dapat dilihat pada formula (3). MAPE hanyalah rata-rata dari persentase kesalahan.

$$MAPE = \frac{1}{n} \sum \frac{|e_t|}{d_t} \dots\dots\dots (3)$$

**Data Penelitian**

Data dikumpulkan dalam format nilai yang dipisahkan koma (CSV). Data harga saham dikumpulkan dalam kurun waktu empat tahun dari tahun 2016 sampai dengan September 2020. Data fundamental dari perusahaan dikumpulkan dari Laporan Keuangan masing-masing perusahaan. Laporan keuangan dirilis setiap triwulan di setiap tahun. Data harga saham dikumpulkan dari Google Finance yang dirilis setiap hari di hari kerja. Detail perusahaan dapat dilihat pada tabel 1.

Tabel 1. Data Penelitian

Nama Perusahaan KODE	
Telekomunikasi Indonesia	TLKM
XL Axiata	EXCL
Smartfren	FREN
Indosat	ISAT

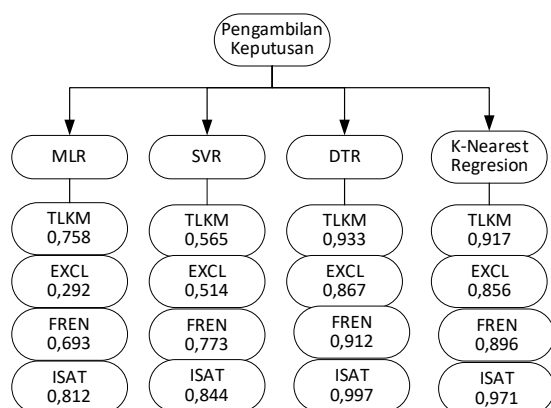
Sumber: [www.idx.co.id](http://www.idx.co.id)

Data yang akan diolah dari perusahaan yang akan digunakan adalah

total aset lancar, total kewajiban, laba bersih periode berjalan, dan *Earning Per Share* (EPS). Aset lancar termasuk aset perusahaan yang dapat dengan mudah dicairkan dalam bentuk uang dalam waktu kurang dari satu tahun seperti kas, surat berharga, piutang, piutang pendapatan, biaya dibayar di muka, perlengkapan, persediaan barang dagangan. Kewajiban adalah jumlah yang terutang atau kewajiban perusahaan untuk membayar kepada pihak lain. EPS adalah tingkat laba bersih untuk setiap saham yang dapat dicapai perusahaan pada saat menjalankan operasinya pada periode berjalan. Perhitungan EPS diperoleh dari laba yang tersedia bagi pemegang saham biasa dibagi dengan jumlah saham biasa.

**PEMBAHASAN**

Hasil penelitian dan pembahasan akan dibahas pada bagian ini. Secara sederhana pada gambar 7 digambarkan pohon keputusan yang didapat dari hasil pengolahan data pada penelitian ini.



Gambar 7. Pohon Keputusan dari hasil pengolahan data

Tabel 2 menunjukkan hasil koefisien korelasi dari model regresi pada masing-masing perusahaan. Secara keseluruhan, ini menyajikan hasil yang baik pada koefisien korelasi (>0,5) kecuali pada EXCEL dengan algoritma Regresi Linier Berganda. Hal ini menunjukkan bahwa meskipun dalam situasi era covid19, data fundamental dari perusahaan masih berkorelasi dengan harga saham. Koefisien korelasi tertinggi adalah dengan Regresi Pohon Keputusan.

Algoritma ini menunjukkan koefisien korelasi yang tinggi.

Tabel 2. Hasil Koefisien Korelasi

	TLKM	EXCL	FREN	ISAT
MLR	0.758	0.292	0.693	0.812
SVR	0.565	0.514	0.773	0.844
DTR	<b>0.933</b>	<b>0.867</b>	<b>0.912</b>	<b>0.997</b>
K-Nearest Reg.	0.917	0.856	0.896	0.971

Sumber: Olah Data

Hasil MAPE dengan data split uji common train dapat dilihat pada tabel 3. Angka tersebut mencapai hasil yang berbeda dengan algoritma yang telah digunakan. Meskipun demikian, regresi pohon keputusan menunjukkan angka yang mengesankan dengan hasil MAPE terkecil sebesar 2,99% dan terbesar sebesar 12,87 %. Algoritma kedua yang menunjukkan hasil yang baik adalah K-Nearest Regression. Hasil MAPE menunjukkan angka yang sedikit berbeda dengan Regresi Pohon Keputusan dan Regresi K-Nearest. Algoritma yang kurang cocok dengan kasus ini adalah regresi linier berganda. Algoritma ini menunjukkan hasil MAPE terkecil sebesar 11,476% dan jumlah maksimum sebesar 73,178% mempengaruhi pasar saham, hasil tersebut menunjukkan bahwa data fundamental dan harga saham memiliki hubungan. Koefisien korelasi yang tinggi dihasilkan dari Regresi Pohon Keputusan dan Regresi K-Nearest. Regresi Pohon Keputusan menghasilkan hasil yang baik pada data latih pada Test Split dan validasi silang K-Fold, 2,99% dan 2,98% secara berulang. Karya masa depan penelitian ini akan menambahkan lebih banyak data dan variabel dan menggunakan analisis sentimen dari berita atau opini publik tentang saham dengan metode yang lebih canggih.

Tabel 3. Hasil MAPE dengan data uji split

	TLKM	EXCL	FREN	ISAT
MLR	11.42 %	16.55 %	73.18 %	54.98 %
SVR	9.27 %	15.82 %	50.77 %	52.22 %
DTR	2.98 %	7.76 %	12.87 %	8.32 %
K-Nearest Reg.	3.23 %	8.04 %	12.98 %	8.79 %

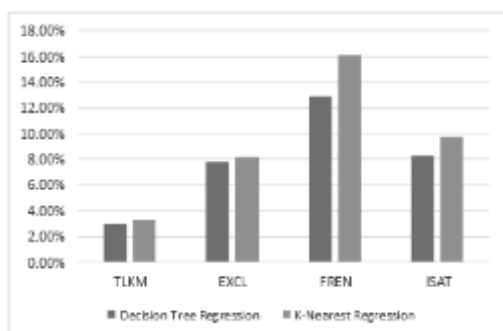
Sumber: Olah Data

K-Fold digunakan kali ini. Hasil MAPE dari data splitting yang berbeda dapat dilihat pada Tabel 4. Hasil penelitian menunjukkan tidak banyak perbedaan dengan data split uji common train. Dua algoritma yang memprediksi jumlah MAPE paling minimum adalah Regresi Pohon Keputusan dan Regresi K-Nearest. Namun, Regresi Linier Berganda menunjukkan hasil yang sedikit lebih baik dengan K-Fold. Hasil MAPE paling sedikit 11,42% dan paling banyak 72,15%. Perbandingan hasil MAPE dari Regresi Pohon Keputusan dapat dilihat pada Gambar 8. Perbedaan hasil MAPE dan KFold hanya sedikit.

Tabel 4. Hasil MAPE dengan K-Fold

	TLKM	EXCL	FREN	ISAT
MLR	11.42 %	16.57 %	72.15 %	54.795 %
SVR	9.27 %	15.86 %	52.31 %	47.59 %
DTR	<b>2.99 %</b>	<b>7.77 %</b>	<b>12.91%</b>	<b>8.28 %</b>
K-Nearest Reg.	3.31 %	8.18 %	16.13 %	9.78 %

Sumber: Olah Data



Gambar 8. Hasil MAPE dengan Regresi Pohon Keputusan

## SIMPULAN

Dalam karya ini, kami memprediksi harga saham telekomunikasi Indonesia termasuk era covid19 dengan Regresi Linier Berganda, Regresi Vektor Pendukung, Regresi Pohon Keputusan, dan Regresi K-Nearest. Data fundamental historis selama 3 tahun digunakan untuk memprediksi harga saham. Meskipun situasi pandemi sangat mempengaruhi pasar saham, namun hasil penelitian menunjukkan bahwa data fundamental dan harga saham memiliki

hubungan. Koefisien korelasi yang tinggi dihasilkan dari Regresi Pohon Keputusan dan *Regresi K-Nearest*. Regresi Pohon Keputusan menghasilkan hasil yang baik pada data *Train Test Split* dan *KFold Cross Validation*, 2,99% dan 2,98% secara berulang. Karya masa depan penelitian ini akan menambahkan lebih banyak data dan variabel dan menggunakan analisis sentimen dari berita atau opini publik tentang saham dengan metode yang lebih canggih.

## REFERENSI

- Agustin, A. H., Gandhiadi, G. K., & Oka, T. B. (2016). *Penerapan Metode Fuzzy Sugeno Untuk Menentukan Harga Jual Sepeda Motor Bekas*. 5(4), 176–182.
- Asghar, M. Z., Rahman, F., Kundi, F. M., & Ahmad, S. (2019). Development of stock market trend prediction system using multiple regression. *Computational and Mathematical Organization Theory*, 25(3), 271–301. <https://doi.org/10.1007/s10588-019-09292-7>
- Boujelbene, S. Z., Mezghani, D. B. A., & Ellouze, N. (2010). Improving svm by modifying kernel functions for speaker identification task. *International Journal of Digital Content Technology and Its Applications*, 4(6), 100–105. <https://doi.org/10.4156/jdcta.vol4.issue.6.12>
- Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., & Li, Z. (2018). A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sensing of Environment*, 210, 35–47. <https://doi.org/10.1016/j.rse.2018.02.045>
- Cakra, Y. E., & Distiawan Trisedya, B. (2015). Stock price prediction using linear regression based on sentiment

- analysis. In *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. IEEE. <https://doi.org/10.1109/icacsis.2015.7415179>
- Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, *80*, 340–355. <https://doi.org/10.1016/j.eswa.2017.02.044>
- Czajkowski, M., & Kretowski, M. (2016). The role of decision tree representation in regression problems – An evolutionary perspective. *Applied Soft Computing*, *48*, 458–475. <https://doi.org/10.1016/j.asoc.2016.07.007>
- Farida, Y. (2016). Sistem Prediksi Saham Menggunakan Adaptive Neuro Fuzzy Inference System (Studi Kasus Saham Mingguan PT Astra Agro Lestari,Tbk). *Systemic: Information System and Informatics Journal*, *2*(2), 35–39. <https://doi.org/10.29080/systemic.v2i2.113>
- Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., & Anastasiu, D. C. (2019). Stock Price Prediction Using News Sentiment Analysis. In *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*. IEEE. <https://doi.org/10.1109/bigdataservice.2019.00035>
- Nguyen, L. A., Verreth, J. A. J., Leemans, R., Bosma, R., & De Silva, S. (2016). A decision tree analysis to support potential climate change adaptations of striped catfish (*Pangasianodon hypophthalmus* Sauvage) farming in the Mekong Delta, Vietnam. *Tropicultura*, *34*(Specialissue), 105–115. <https://library.wur.nl/WebQuery/wurpubs/512763>
- Olive, D. J. (2017). Multiple Linear Regression. In *Linear Regression* (pp. 17–83). Springer International Publishing. [https://doi.org/10.1007/978-3-319-55252-1\\_2](https://doi.org/10.1007/978-3-319-55252-1_2)
- Ouahilal, M., Mohajir, M. El, Chahhou, M., & Mohajir, B. E. El. (2017). A novel hybrid model based on Hodrick–Prescott filter and support vector regression algorithm for optimizing stock market price prediction. *Journal of Big Data*, *4*(1). <https://doi.org/10.1186/s40537-017-0092-5>
- Parray, I. R., Khurana, S. S., Kumar, M., & Altalbe, A. A. (2020). Time series data analysis of stock price movement using machine learning techniques. *Soft Computing*, *24*(21), 16509–16517. <https://doi.org/10.1007/s00500-020-04957-x>
- Purnama, R. B. (2017). Perancangan Prediksi Untuk Menentukan Indeks Harga Saham Menggunakan Jaringan Syaraf Tiruan. *Kinetik*, *2*(2), 125. <https://doi.org/10.22219/kinetik.v2i2.190>
- Rifa'i, A., & Mahdiana, D. (2019). Optimized Fuzzy Backpropagation Neural Network using Genetic Algorithm for Predicting Indonesian Stock Exchange Composite Index. In *2019 6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. IEEE. <https://doi.org/10.23919/eecsi48112.2019.8977058>
- Rustam, Z., Vibranti, D. F., & Widya, D. (2018). *Predicting the direction of Indonesian stock price movement using support vector machines and fuzzy Kernel C-Means*. Author(s). <https://doi.org/10.1063/1.5064205>
- Rustam, Zuherman, & Kintandani, P.



- (2019). Application of Support Vector Regression in Indonesian Stock Price Prediction with Feature Selection Using Particle Swarm Optimisation. *Modelling and Simulation in Engineering*, 2019, 1–5. <https://doi.org/10.1155/2019/8962717>
- Selvamuthu, D., Kumar, V., & Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, 5(1). <https://doi.org/10.1186/s40854-019-0131-7>
- Siew, H. L., & Nordin, M. J. (2012). Regression techniques for the prediction of stock price trend. In *2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE)*. IEEE. <https://doi.org/10.1109/icssbe.2012.6396535>
- Vanneschi, L., Castelli, M., Manzoni, L., Silva, S., & Trujillo, L. (2020). Is k Nearest Neighbours Regression Better Than GP? In *Lecture Notes in Computer Science* (pp. 244–261). Springer International Publishing. [https://doi.org/10.1007/978-3-030-44094-7\\_16](https://doi.org/10.1007/978-3-030-44094-7_16)
- Xianya, J., Mo, H., & Haifeng, L. (2019). Stock Classification Prediction Based on Spark. *Procedia Computer Science*, 162(Itqm), 243–250. <https://doi.org/10.1016/j.procs.2019.11.281>
- Yu, H., Chen, R., & Zhang, G. (2014). A SVM Stock Selection Model within PCA. *Procedia Computer Science*, 31, 406–412. <https://doi.org/10.1016/j.procs.2014.05.284>
- Zhang, D., Hu, M., & Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, 36, 101528. <https://doi.org/10.1016/j.frl.2020.101528>
- Zhang, G., Xu, L., & Xue, Y. (2017). Model and forecast stock market behavior integrating investor sentiment analysis and transaction data. *Cluster Computing*, 20(1), 789–803. <https://doi.org/10.1007/s10586-017-0803-x>