



Impurity-Based Important Features for feature selection in Recursive Feature Elimination for Stock Price Forecasting

Arif Mudi Priyatno^{1✉}, Wahyu Febri Ramadhan Sudirman², R. Joko Musridho³, Fazilla Amalia⁴

Program Studi Bisnis Digital, Universitas Pahlawan Tuanku Tambusai^{(1),(4)}

Program Studi Perbankan Islam, Universitas Pahlawan Tuanku Tambusai⁽²⁾

Program Studi Teknik Informatika, Universitas Pahlawan Tuanku Tambusai⁽³⁾

DOI: 10.31004/jutin.v6i4.17726

✉ Corresponding author:

[arifmudi@universitaspahlawan.ac.id]

Article Info	Abstrak
<p><i>Kata kunci:</i> <i>Fitur Penting;</i> <i>Impurity;</i> <i>Recursive Feature Elimination;</i> <i>Various Moving Average;</i> <i>Machine Learning</i></p>	<p>Investor saham melakukan peramalan harga saham berdasarkan teknikal indikator dan histori harga saham. Jumlah teknikal indikator dan data histori yang tidak sedikit membuat machine learning mengalami overfitting dan ambiguitas dalam melakukan peramalan. Penelitian ini mengusulkan seleksi fitur menggunakan fitur penting berbasis impurity pada <i>recursive feature elimination</i> untuk peramalan harga saham. Data digunakan adalah data histori dan data <i>various moving average</i>. Seleksi fitur digunakan untuk mengurangi jumlah fitur dan mendapatkan fitur yang penting dan relevan. Metode seleksi fitur yang digunakan adalah <i>recursive feature elimination</i> dengan fitur penting berbasis impurity. Metode <i>machine learning</i> yang digunakan <i>linear regression</i>, <i>support vector regression</i>, <i>multi-layer perceptron regression</i>, dan <i>random forest regression</i>. Hasil pengukuran mse, rmse, mae, dan mape fitur optimal paling baik didapatkan dengan jumlah 6 fitur dan <i>machine learning linear regression</i>. Hasil rata-rata mse, rmse, mae, dan mape yaitu 0.000279, 0.016577, 0.012843, dan 1.42236%. Hal ini membuktikan fitur penting berbasis impurity untuk seleksi fitur pada <i>recursive feature elimination</i> dengan menggunakan data histori dan <i>various moving average</i> berhasil melakukan peramalan.</p>
<p><i>Keywords:</i> <i>Important Features;</i> <i>Impurity;</i> <i>Recursive Feature Elimination;</i> <i>Various Moving Averages;</i> <i>Machine Learning</i></p>	<p>Abstract</p> <p>Stock investors perform stock price forecasting based on technical indicators and historical stock prices. The large number of technical indicators and historical data often leads to overfitting and ambiguity in forecasting using machine learning. In this paper, we proposed a feature selection approach using impurity-based important features in recursive feature elimination for stock price forecasting. The data utilized includes historical data and various moving averages. Feature selection is employed to reduce the number of features and obtain important and</p>

relevant features. The recursive feature elimination with impurity-based important features is utilized as the feature selection method. The machine learning methods employed are linear regression, support vector regression, multi-layer perceptron regression, and random forest regression. The measurement results of mean squared error (mse), root mean squared error (rmse), mean absolute error (mae), and mean absolute percentage error (mape) show that the optimal feature selection and machine learning method is achieved with six features and linear regression. The average mse, rmse, mae, and mape values are 0.000279, 0.016577, 0.012843, and 1.42236%, respectively. These results validate the effectiveness of impurity-based important features for feature selection in recursive feature elimination using historical data and various moving averages in stock price forecasting.

1. INTRODUCTION

Peramalan harga saham memiliki peran yang sangat penting dalam dunia investasi dan perdagangan saham (Ji et al., 2021). Saham sebagai instrumen keuangan memberikan kepemilikan kepada pemegang saham dalam suatu perusahaan (Charan et al., 2022), dan pasar saham menjadi tempat perdagangan saham antara investor dan perusahaan (Kumar et al., 2021). Dalam aktivitas perdagangan saham, investor mengandalkan indikator teknikal untuk menganalisis kinerja perusahaan dan meramalkan pergerakan harga saham (Lee et al., 2022).

Pentingnya peramalan harga saham terletak pada kebutuhan investor dan pelaku pasar untuk mengambil keputusan investasi yang tepat (Hamzah et al., 2021). Dengan memiliki peramalan yang akurat, investor dapat mengoptimalkan keuntungannya dan mengurangi risiko yang terkait dengan perdagangan saham (Park et al., 2022). Peramalan harga saham yang baik dapat membantu investor dalam melakukan pembelian dan penjualan saham secara tepat waktu, sehingga dapat memanfaatkan peluang pasar dan menghindari potensi kerugian (Bouktif et al., 2019; Kehinde et al., 2023; Namdari & Durrani, 2021).

Kemajuan teknologi informasi dan kecerdasan buatan saat ini membantu penemu untuk meramalkan harga saham (Chandrika & Srinivasan, 2021; Ray et al., 2018). *Machine learning* adalah bagian dari kecerdasan buatan yang dapat digunakan untuk peramalan harga saham. Metode machine learning yang dapat digunakan untuk peramalan harga saham antara lain *multi-layer perceptron* (Azizah et al., 2020; Namdari & Durrani, 2021; Sundar & Satyanarayana, 2019), *linear regression* (Chauhan & Sharma, 2023; Deshmukh et al., 2023; Meghana & Arumugam, 2023), *support vector regression* (Azimifar et al., 2020; Fang, 2020; Huang et al., 2022), dan *random forest regression* (Ghosh et al., 2022; Illa et al., 2022; Lavingia et al., 2022).

Penelitian Xia dkk (2013) memprediksi tren saham. Metode *machine learning* digunakan pada penelitian ini adalah *support vector regression*. Data digunakan pada penelitian ini ada histori data saham. Histori tersebut terdiri atas *open*, *high*, *low*, *volume*, dan *adj_close*. Penelitian Bhuriya dkk (2017) memprediksi saham menggunakan *machine learning linear regression*. Data digunakan yaitu *open price*, *high price*, *low price*, dan *close price*. Hasil penelitian menunjukkan bahwa *machine learning linear regression* lebih baik di bandingkan *radial basic function*. Penelitian ini tidak membahas secara mendalam tentang hasil penelitian dan fitur data masih terbatas pada data histori. Penelitian Vijn dkk (2020) memprediksi harga *close* saham menggunakan teknik *machine learning*. Data digunakan yaitu *high*, *low*, *open*, *close*, *adjacent close*, dan *volume*. Metode *machine learning* digunakan yaitu *artificial neural network*, dan *random forest regression*.

Peramalan harga saham menggunakan data transaksi harian historis dan indikator teknikal (Albahli et al., 2022). Data histori transaksi harian yang umum digunakan seperti *open*, *close*, *high*, *low*, dan *volume* (K et al., 2021). Data teknikal indikator seperti *simple moving average*, *relative strength index*, dan lain-lain (Basak et al., 2019; Pandya & Jaliya, 2022; Park et al., 2022), Jumlah fitur yang dihasilkan dari data riwayat transaksi harian dan teknikal indikator sangat banyak. Hal ini menyebabkan ambiguitas dan overfitting dalam peramalan harga saham menggunakan *machine learning* dan mengurangi kinerja peramalan.

Dalam penelitian ini, kami mengusulkan seleksi fitur menggunakan fitur penting berbasis impurity pada *recursive feature elimination* untuk peramalan harga saham. Seleksi fitur digunakan untuk mengurangi jumlah fitur dan mendapatkan fitur yang penting dan relevan. Metode seleksi fitur yang digunakan adalah *recursive feature elimination* dengan fitur penting berbasis impurity. *Machine learning* digunakan untuk pemodelan peramalan harga saham. Metode *machine learning* yang digunakan *linear regression*, *support vector regression*, *multi-layer perceptron regression*, dan *random forest regression*. Penelitian ini dibandingkan dengan penelitian (Bhuriya et al., 2017; Vijn et al., 2020; Xia et al., 2013).

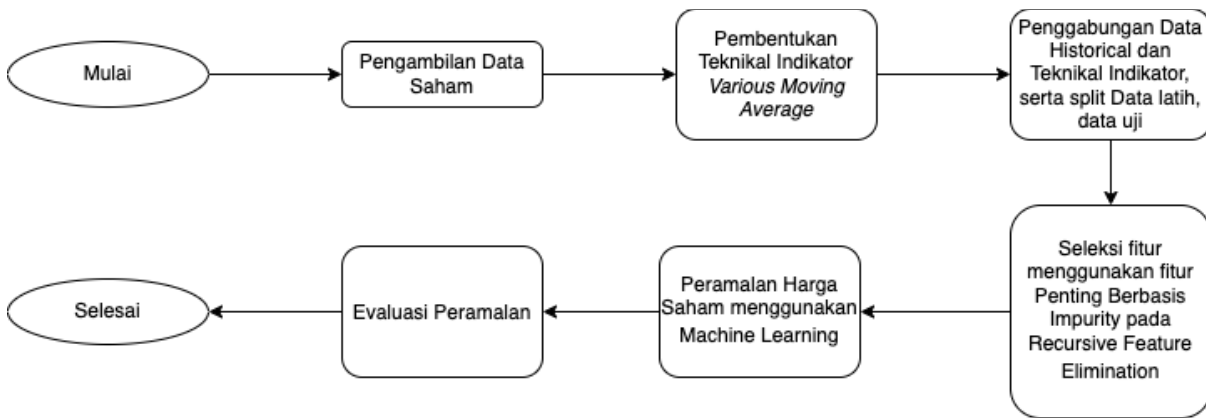


Figure 1 Tahapan Penelitian Seleksi Fitur untuk Peramalan Harga Saham

2. METHODS

Proses penelitian ini diantaranya yaitu pengambilan data, pembentukan data teknikal indikator, penggabungan data historical dan data teknikal indikator serta split/ pembagian data latih data uji, seleksi fitur menggunakan fitur penting berbasis impurity, peramalan harga saham menggunakan machine learning, dan evaluasi kinerja peramalan. Gambar 1 menunjukkan proses dari penelitian ini.

2.1 PENGAMBILAN DATA SAHAM

Dataset saham PT Bank Central Asia Tbk (BBCA) diambil dari yahoo finance. Dataset diambil dari 20 April 2015 hingga 24 Maret 2023. Fitur dataset adalah *previous, close, open price, first trade, high, low, offer, bid, volume, value, offer volume, bid volume, foreign sell, foreign buy, non-regular volume*, dan *non-regular value*. Saham BBCA melakukan stock split pada 12 Oktober 2021 dengan faktor split 5 (1:5). Gambar 2 merupakan grafik dari history data saham PT Bank Central Asia Tbk (BBCA).

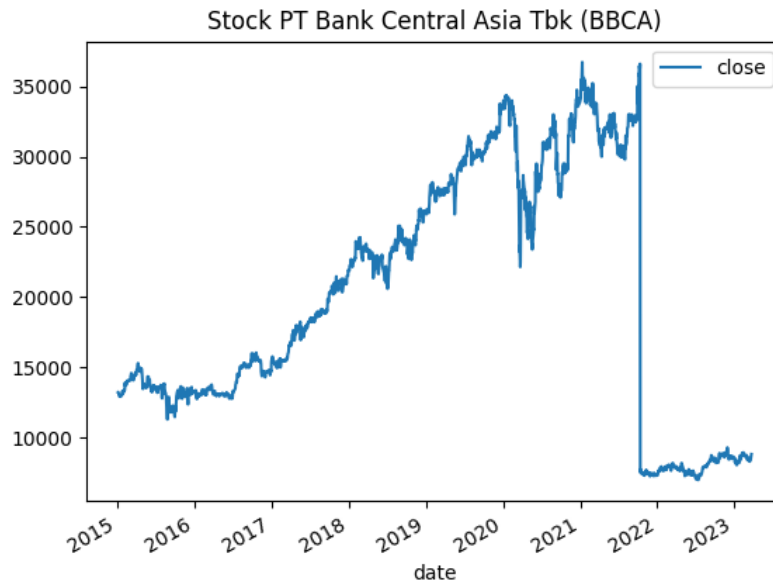


Figure 2 Grafik Harga Saham BBCA berserta stock split 12 oktober 2021

2.2 PEMBENTUKAN TEKNIKAL INDIKATOR DENGAN VARIOUS MOVING AVERAGE

Various moving average digunakan pada penelitian ini yaitu *simple moving average (SMA)*, *exponential moving average (EMA)*, dan *real moving average (RMA)*. *Moving average* merupakan salah satu indikator teknikal analisis yang digunakan untuk memperhalus data pergerakan harga dan membantu trader dalam melakukan identifikasi tren harga saham. *Simple moving average* merupakan *moving average* paling sederhana. *Simple moving average* dihitung dengan cara menjumlahkan deretan sebanyak periode dan di bagi dengan periode yang ditentukan tersebut. Persamaan 1 merupakan cara menghitung *simple moving average* (Ma et al., 2023; Yeo et al., 2023). *Exponential Moving Average (EMA)* merupakan jenis *moving average* yang memberikan bobot lebih kepada harga terbaru dalam menghitung rata-rata. EMA merespon perubahan harga lebih cepat dibanding dengan SMA, karena memberikan bobot pada harga terbaru. hal ini membuat EMA lebih cepat respon terhadap harga dan kadang memberikan signal palsu. Persamaan 2 merupakan cara menghitung

Exponential moving average (Ayyappa & Siva Kumar, 2023; Ma et al., 2023; Yeo et al., 2023). *Real moving average* merupakan jenis perhitungan moving average yang memberikan bobot yang lebih besar pada harga terakhir, sehingga memberikan sinyal lebih cepat dibandingkan dengan jenis moving average lainnya. Persamaan 3 merupakan cara menghitung *real moving average* (Yeo et al., 2023).

$$SMA_t(CP, n) = \frac{1}{n} \sum_{k=t-n+1}^t CP_k \tag{1}$$

$$EMA_t(CP, n) = \alpha \times CP_t + (1 - \alpha) \times EMA_{t-1}(n - 1) \tag{2}$$

$$RMA_t(CP, n) = (CP_t - RMA_{t-1}) \times \frac{n-1}{2} + RMA_{t-1} \tag{3}$$

Dimana CP merupakan *close price*, dan n merupakan periode moving average yang ditentukan.

2.3 PENGABUNGAN DATA, NORMALISASI, DAN SPLIT DATA LATIH - DATA UJI

Data histori harian harga saham digabungkan dengan data teknikal indikator *various moving average*. Penggabungan dilakukan dengan menyatukan dataset histori dan data teknikal indikator berdasarkan index harian transaksi. Data hasil penggabungan dilakukan proses normalisasi min-max. Persamaan 4 merupakan cara untuk melakukan normalisasi min-max (Ampomah et al., 2021). Data hasil normalisasi dilakukan pembagian data latih dan data uji. Split data latih dan data uji dilakukan dengan cross-validation. Gambar 3 merupakan proses cross-validation dan pembagian data latih – data uji.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

Dimana x adalah data asli, x_min adalah nilai minimum dari data, n_max adalah nilai maksimum dari data. Dan x_norm merupakan data hasil normalisasi.

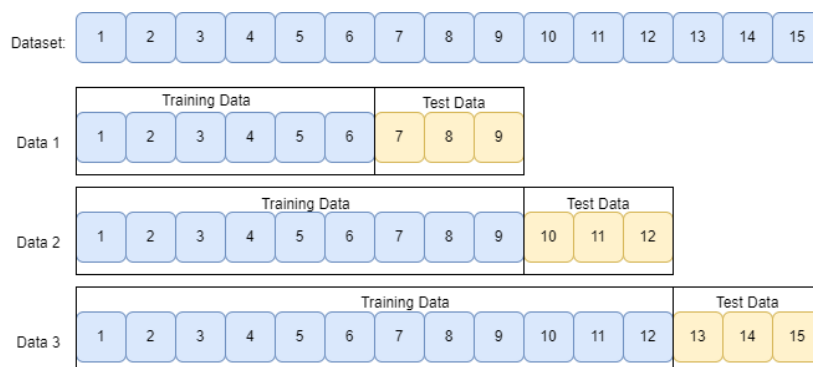


Figure 3 Strategi Split data latih dan data uji menggunakan cross-validation

2.4 SELEKSI FITUR MENGGUNAKAN FITUR PENTING BERBASIS IMPURITY PADA RECURSIVE FEATURE ELIMINATION

Fitur penting berbasis impurity didapatkan dari fitur penting bawaan random forest regression. Impurity bawaan merujuk pada ukuran ketidakmurnian atau ketidaksempurnaan dalam set data pelatihan yang digunakan oleh setiap pohon keputusan dalam random forest. Dalam random forest regression, setiap pohon mempertimbangkan impurity bawaan untuk membagi set data menjadi subset yang lebih kecil. Pohon-pohon ini kemudian menggabungkan hasil prediksi mereka secara kolektif untuk menghasilkan prediksi akhir. Dengan mempertimbangkan impurity bawaan, random forest regression dapat memilih fitur-fitur yang paling informatif dan relevan untuk memprediksi variabel target. Ini membantu mengurangi overfitting dan meningkatkan kualitas prediksi model secara keseluruhan. Dengan menggunakan impurity bawaan sebagai landasan, random forest regression memungkinkan model untuk mengambil keputusan yang lebih baik dan memberikan hasil yang lebih akurat dalam tugas regresi.

Algoritma 1 Seleksi fitur berbasis impurity pada recursive feature elimination

Input : Fitur Hasil penggabungan fitur
Jumlah fitur yang diinginkan

Parameter <i>random forest regressor</i>	
Output	: Daftar Fitur Relevan untuk prediksi harga saham
Step	: 1. Masukkan semua fitur hasil penggabungan fitur
	2. Masukkan jumlah fitur yang diinginkan
	3. Masukkan parameter <i>random forest regressor</i>
	4. Proses Pelatihan menggunakan <i>random forest regressor</i>
	5. Hitung Fitur Penting berbasis impurity hasil pelatihan <i>random forest regressor</i>
	6. Urutkan Fitur Penting Berbasis Impurity dari tertinggi hingga terendah
	7. Lakukan Eliminasi terhadap fitur dengan pengaruh terendah
	8. Ulangi langkah 4 sampai 7 hingga semua fitur telah tereliminasi dan didapatkan urutan pengaruh fitur
	9. Dapatkan fitur berpengaruh sesuai dengan jumlah yang diinginkan
	10. Fitur relevan yang berpengaruh untuk prediksi harga saham disimpan pada matrik.

Fitur penting berbasis impurity hasil dari *random forest regression* ini digunakan sebagai pertimbangan tingkat pentingnya fitur pada proses *recursive feature elimination*. Algoritma 1 merupakan langkah dari algoritma seleksi fitur berbasis impurity pada *recursive feature elimination*.

2.5 PERAMALAN HARGA SAHAM MENGGUNAKAN MACHINE LEARNING

Machine learning untuk peramalan harga saham pada penelitian ini yaitu *support vector regression*, *linear regression*, *random forest regression*, dan *multi-layer perceptron regression*. *Machine learning* metode *linear regression* menggunakan penelitian (Bhuriya et al., 2017), metode *multi-layer perceptron regression* menggunakan penelitian (Vijh et al., 2020), metode *random forest regression* menggunakan penelitian (Vijh et al., 2020), dan metode *support vector regression* menggunakan penelitian (Xia et al., 2013).

2.6 EVALUASI PERAMALAN

Evaluasi kinerja dari penelitian ini menggunakan Mean Square Error (MSE) ditunjukkan pada Persamaan 5, Root Mean Square Error (RMSE) ditunjukkan pada Persamaan 6, Mean Absolute Error (MAE) ditunjukkan pada Persamaan 7, dan Mean Percentage Absolute Error (MAPE) ditunjukkan pada Persamaan 8.

$$mse = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{5}$$

$$rmse = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{6}$$

$$mae = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{7}$$

$$mape = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \tag{8}$$

Dimana n merupakan jumlah total data, y_i merupakan nilai target asli, dan \hat{y}_i merupakan hasil prediksi target yang didapatkan. Penelitian ini dibandingkan dengan paper (Bhuriya et al., 2017; Vijh et al., 2020; Xia et al., 2013).

3. RESULT AND DISCUSSION

Jumlah dataset historical saham BBCA didapatkan yaitu 1931 row data. Dataset perlu dilakukan penyesuaian dengan aksi korporasi stock split BBCA dengan perbandingan 1: 5 (lama:baru) pada tanggal 12 oktober 2021. Gambar 4 merupakan grafik harga saham BBCA setelah disesuaikan dengan aksi stock split.

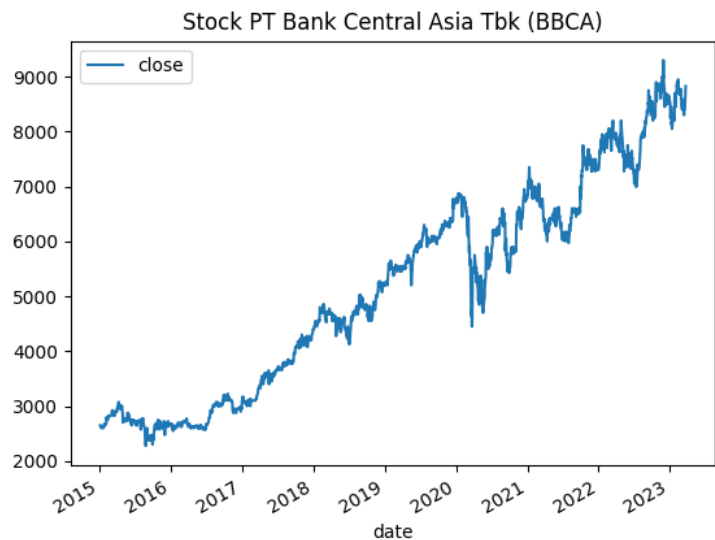


Figure 4 Grafik harga BBCA setelah penyesuaian stock split

Pembentukan teknikal indikator *various moving average* diantaranya yaitu *simple moving average*, *exponential moving average*, dan *real moving average*. Periode pembentukan *various moving average* yaitu 3 sampai 75. Total fitur teknikal indikator *various moving average* terbentuk yaitu 216 fitur. Tabel 2 merupakan sample hasil dari proses pembentukan *various moving average*. Sample hasil teknikal indikator pada Tabel 2 yaitu SMA_70, SMA_75, EMA_70, EMA_75, RMA_70, dan RMA_75 pada 20 april 2015 diantaranya 2824, 2811, 2856, 2849, 2850, dan 2847.

Table 1 Sample data teknikal indikator various moving average

Date	SMA_70	SMA_75	EMA_70	EMA_75	RMA_70	RMA_75
2015-04-21	2826	2813	2859	2851	2853	2850
2015-04-22	2831	2817	2862	2854	2855	2853
2015-04-23	2837	2822	2866	2858	2858	2856
2015-04-24	2843	2827	2870	2862	2861	2859
2015-04-27	2846	2830	2868	2861	2861	2858

Data historical dan data teknikal indikator various moving average dilakukan penggabungan. Jumlah fitur hasil penggabungan data historical dan data teknikal indikator yaitu 235 fitur. Data hasil penggabungan dilakukan proses pembagian data latih dan data uji menggunakan cross-validation. Tabel 3 merupakan hasil dari cross-validation data latih dan data uji. Cross-validation digunakan yaitu 5 CV. Jumlah data test digunakan adalah 20 data. Cross-validation dari 1 hingga 5 secara berurutan yaitu 1831, 1851, 1871, 1891, dan 1911 data.

Table 2 Cross-validation data latih dan data uji

Cross-Validation	Number of Train Data	Number of Test Data
1	1831	20
2	1851	20
3	1871	20
4	1891	20
5	1911	20

Data hasil pembagian data latih dan data uji dilakukan proses seleksi fitur. Seleksi fitur berdasarkan fitur penting berbasis impurity menggunakan *recursive feature elimination*. Pengaturan seleksi fitur digunakan yaitu 1 hingga 235 fitur. Tabel 4 merupakan hasil seleksi fitur dengan n fitur 5, 6, dan 7 fitur.

Table 3 Fitur hasil seleksi dengan n 5, 6, dan 7 fitur

n fitur	CV 1	CV 2	CV 3	CV 4	CV 5
---------	------	------	------	------	------

5	close, ema_3, bid, offer, sma_75	close, sma_75, offer, ema_3, bid	close, sma_3, bid, offer, sma_75	close, bid, offer, sma_3, sma_74	close, offer, ema_3, bid, sma_73
6	close, ema_3, bid, offer, sma_75, ema_5	close, sma_75, offer, ema_3, bid, sma_66	close, sma_3, bid, offer, sma_75, ema_3	close, bid, offer, sma_3, ema_3, sma_74	close, offer, sma_3, bid, sma_73, ema_3
7	close, ema_3, bid, offer, sma_75, ema_5, sma_72	close, sma_75, offer, ema_3, bid, sma_66, sma_3	close, sma_3, bid, offer, sma_75, ema_3, sma_73	close, bid, offer, sma_3, ema_3, sma_74, sma_75	close, offer, sma_3, bid, sma_73, ema_3, sma_42

Tabel 4 menunjukkan bahwa fitur yang penting berbasis impurity dengan jumlah n fitur 5 CV1 (close, ema_3, bid, offer, sma_75), CV2 (close, sma_75, offer, ema_3, bid), CV3 (close, sma_3, bid, offer, sma_75), CV4 (close, bid, offer, sma_3, sma_74), dan CV5 (close, offer, ema_3, bid, sma_73). Hasil cross-validation sebanyak 5 kali pada jumlah fitur 5 yang selalu muncul yaitu close, bid, dan over. Fitur lainnya berubah-ubah sesuai dengan jumlah data latih yang digunakannya. Hal ini menunjukkan histori jumlah data yang digunakan mempengaruhi terhadap teknikal indikator yang penting pada peramalannya. Fitur penting ini terus ada pada jumlah fitur lainnya. Hal ini menunjukkan bahwa close / penutupan harga, bid / permintaan, offer / transaksi berhasil, *simple moving average*, dan *exponential moving average* memiliki nilai penting untuk peramalan harga saham. Fitur penting Tabel 4 dijadikan sebagai fitur imputasi pada proses pelatihan dan peramalan harga saham menggunakan machine learning. Machine learning digunakan yaitu *linear regression*, *multi-layer perceptron regression*, *random forest regression*, dan *support vector regression*.

Table 4 peramalan menggunakan seleksi fitur dan linear regression

<i>Mean absolute error (MAE)</i>						
n fitur	CV1	CV2	CV3	CV4	CV5	Means
6	0.011311	0.015555	0.011414	0.01334	0.012596	0.012843
7	0.011303	0.015479	0.01134	0.013536	0.012649	0.012862
5	0.011392	0.015546	0.01153	0.013251	0.012663	0.012876
<i>Mean absolute percentage error (MAPE)</i>						
6	1.21649%	1.68688%	1.33563%	1.46387%	1.40894%	1.42236%
7	1.21513%	1.67844%	1.32637%	1.48550%	1.41509%	1.42411%
5	1.22517%	1.68590%	1.34907%	1.45423%	1.41646%	1.42616%
<i>Mean squared error (MSE)</i>						
6	0.000175	0.000377	0.000234	0.000321	0.00029	0.000279
5	0.000176	0.000379	0.000233	0.000317	0.000292	0.000279
12	0.00017	0.000352	0.000255	0.00034	0.000292	0.000282
<i>Root Mean Squared Error (RMSE)</i>						
6	0.013217	0.019425	0.015298	0.017905	0.017038	0.016577
5	0.01328	0.019469	0.01527	0.017798	0.017091	0.016582
7	0.013251	0.019395	0.015238	0.018205	0.017165	0.016651

Tabel 5 menunjukkan hasil terbaik dari peramalan harga saham BBCA menggunakan seleksi future dan *machine learning linear regression*. Hasil menunjukkan jumlah fitur optimal yang selalu muncul berdasarkan *mean absolute error*, *mean absolute percentage error*, *mean squared error*, dan *root mean squared error* yaitu 6 dan 5 fitur. Hasil rata-rata untuk 6 fitur pada *mae*, *mape*, *mse*, dan *rmse* yaitu 0.012843, 1.42236%, 0.000279, dan 0.016577. Hasil rata-rata pada 5 fitur pada *mae*, *mape*, *mse*, dan *rmse* yaitu 0.012876, 1.42616%, 0.000279, dan 0.016582. Jumlah fitur optimal selain 6 dan 5 pada pengukuran *mean absolute error*, *mean absolute percentage error* dan *root mean squared error* yaitu 7 fitur, sedangkan pada pengukuran *mean absolute error* yaitu 12 fitur.

Table 5 Peramalan menggunakan seleksi fitur dan multi-layer perceptron regressor

<i>Mean absolute error (MAE)</i>						
n fitur	CV1	CV2	CV3	CV4	CV5	Means
6	0.010571	0.01153	0.015534	0.014924	0.015689	0.013649
5	0.010143	0.011989	0.012285	0.021639	0.012955	0.013802
203	0.010072	0.012652	0.015736	0.01484	0.015805	0.013821
<i>Mean absolute percentage error (MAPE)</i>						
6	1.131720%	1.246523%	1.829095%	1.629099%	1.776142%	1.522516%

5	1.093316%	1.292557%	1.444330%	2.349643%	1.455173%	1.527004%
203	1.083463%	1.372407%	1.853917%	1.619099%	1.780863%	1.541950%
Mean squared error (MSE)						
6	0.000168	0.000345	0.000381	0.000304	0.000388	0.000317
58	0.000183	0.000475	0.000296	0.000292	0.000357	0.00032
203	0.000159	0.000347	0.000429	0.000265	0.000424	0.000325
Root Mean Squared Error (RMSE)						
6	0.012956	0.018583	0.019529	0.017438	0.01971	0.017643
58	0.01351	0.021785	0.017198	0.017089	0.018882	0.017693
203	0.012613	0.018618	0.02071	0.016265	0.020596	0.01776

Tabel 6 merupakan hasil peramalan menggunakan seleksi fitur dan *machine learning multi-layer perceptron regressor*. Jumlah fitur optimal yang muncul bersamaan pada *mean absolute error*, *mean absolute percentage error*, *mean squared error*, dan *root mean squared error* adalah 6 dan 203 fitur. Hasil rata-rata pengukuran peramalan berdasarkan *mean absolute error*, *mean absolute percentage error*, *mean squared error*, dan *root mean squared error* dengan 6 fitur yaitu 0.013649, 1.522516%, 0.000317, dan 0.017643. Hasil rata-rata pengukuran peramalan menggunakan 203 fitur yaitu 0.013821, 1.541950%, 0.000325, dan 0.01776. Jumlah fitur optimal selain 5 dan 203 berdasarkan pengukuran *mean absolute error*, dan *mean absolute percentage error* yaitu 5 fitur. Hasil rata-rata 5 fitur tersebut yaitu 0.013802, dan 1.527004%. Sedangkan jumlah fitur optimal berdasarkan *mean squared error*, dan *root mean squared error* yaitu 58 fitur. Hasil rata-rata 58 fitur tersebut yaitu 0.00032, dan 0.017693.

Table 6 peramalan menggunakan seleksi fitur dan random forest regressor

Mean absolute error (MAE)						
n fitur	CV1	CV2	CV3	CV4	CV5	Means
1	0.021663	0.0135	0.013841	0.018029	0.0175	0.016906
2	0.02401	0.017491	0.013706	0.017555	0.013783	0.017309
84	0.014772	0.028454	0.02137	0.01447	0.014845	0.018782
Mean absolute percentage error (MAPE)						
1	2.304548%	1.479749%	1.622579%	1.976972%	1.952618%	1.867293%
2	2.554530%	1.923988%	1.610011%	1.919189%	1.537683%	1.909080%
84	1.567447%	3.140961%	2.520094%	1.583412%	1.665163%	2.095415%
Mean squared error (MSE)						
1	0.000667	0.000338	0.00034	0.000498	0.000453	0.000459
2	0.000813	0.000497	0.000357	0.000471	0.000308	0.000489
84	0.000387	0.001024	0.000692	0.000311	0.000349	0.000553
Root Mean Squared Error (RMSE)						
1	0.025835	0.018379	0.018433	0.022318	0.021276	0.021248
2	0.028518	0.022287	0.018893	0.021692	0.017547	0.021788
84	0.019675	0.031993	0.026305	0.017634	0.018694	0.02286

Tabel 7 merupakan hasil peramalan menggunakan seleksi fitur dan *machine learning random forest regressor*. Jumlah fitur optimal yang muncul bersamaan berdasarkan pengukuran *mean absolute error*, *mean absolute percentage error*, *mean squared error*, dan *root squared error* yaitu 1, 2, dan 84 fitur. Hasil rata-rata pengukuran berdasarkan pada jumlah 1 fitur yaitu 0.016906, 1.867293%, 0.000459, dan 0.021248. Hasil rata-rata berdasarkan pada jumlah 2 fitur yaitu 0.017309, 1.909080%, 0.000489, dan 0.021788. Hasil rata-rata berdasarkan pada jumlah 84 fitur yaitu 0.018782, 2.095415%, 0.000553, dan 0.02286. Jarak perbedaan antar jumlah fitur optimal berdasarkan *mean absolute percentage error* adalah 0.1% serta ketiganya dibawah 10%. Hal ini menunjukkan bahwa hasil peramalan berdasarkan *mean absolute percentage error* sangat baik.

Table 7 peramalan menggunakan seleksi fitur dan Support vector regressor

Mean absolute error (MAE)						
n fitur	CV1	CV2	CV3	CV4	CV5	Means
5	0.097997	0.056449	0.014296	0.027853	0.015241	0.042367
6	0.098876	0.055603	0.013606	0.028465	0.015487	0.042407
4	0.098072	0.05564	0.016077	0.027446	0.015487	0.042544
Mean absolute percentage error (MAPE)						
5	10.472490%	6.121854%	1.675896%	3.009002%	1.692599%	4.594368%
6	10.566763%	6.027424%	1.592228%	3.075740%	1.718050%	4.596041%

4	10.480370%	6.031937%	1.890863%	2.964193%	1.722999%	4.618072%
Mean squared error (MSE)						
2	0.009073	0.004343	0.000237	0.001118	0.00045	0.003044
5	0.009936	0.003868	0.000296	0.000955	0.000359	0.003083
4	0.009956	0.003791	0.000382	0.000933	0.000357	0.003084
Root Mean Squared Error (RMSE)						
5	0.099682	0.062193	0.017209	0.030902	0.018953	0.045788
6	0.100528	0.061587	0.016432	0.031654	0.019576	0.045955
4	0.099781	0.061575	0.019545	0.030543	0.018891	0.046067

Tabel 8 merupakan hasil peramalan menggunakan seleksi fitur dan *machine learning support vector regressor*. Jumlah fitur optimal berdasarkan pengukuran *mean absolute error*, *mean absolute percentage error*, *mean squared error*, dan *root mean squared error* yaitu 5 dan 4 fitur. Hasil rata-rata pengukuran menggunakan 5 fitur yaitu 0.042367, 4.594368%, 0.003044, dan 0.045788. Hasil rata-rata pengukuran menggunakan 4 fitur yaitu 0.042544, 4.618072%, 0.003084, dan 0.046067. Jumlah fitur optimal lainnya yang muncul pada pengukuran *mean absolute error*, *mean absolute percentage error*, dan *root squared error* adalah 6 fitur. Hasil rata-rata pengukuran menggunakan 6 fitur yaitu 0.042407, 4.596041%, dan 0.045955. jumlah fitur optimal lainnya berdasarkan pengukuran *mean squared error* yaitu 2 fitur. Hasil rata-rata pengukuran menggunakan 2 fitur adalah 0.003044.

Table 8 Peramalan menggunakan Linear Regression (Bhuriya et al., 2017)

Metrik	CV1	CV2	CV3	CV4	CV5	Means
MSE	0.0002052	0.0003847	0.0002565	0.0004068	0.0003237	0.0003154
RMSE	0.0143242	0.0196125	0.016016	0.0201702	0.0179923	0.017623
MAE	0.0118409	0.0156188	0.0120237	0.0152298	0.0127835	0.0134993
MAPE	1.270862%	1.693128%	1.405889%	1.670003%	1.428955%	1.493767%

Tabel 9 merupakan peramalan menggunakan *machine learning linear regression* dan tanpa seleksi fitur. Hasil rata-rata pengukuran menggunakan *mse*, *rmse*, *mae*, dan *mape* yaitu 0.0003154, 0.017623, 0.0134993, dan 1.493767%. Hasil pengukuran *mean squared error*, *root mean absolute error*, *mean absolute error*, dan *mean absolute percentage error* pada Tabel 9 tanpa teknikal indikator *various moving average* dan *seleksi fitur* dibandingkan dengan metode usulan Tabel 9 menggunakan teknikal indikator *various moving average* dan seleksi fitur menunjukkan bahwa metode usulan lebih unggul. Hasil rata-rata *mean squared error* Tabel 9 yaitu 0.0003154, sedangkan metode usulan Tabel 5 untuk fitur optimal keiganya yaitu 0.000279, 0.000279, dan 0.000282. Hasil rata-rata *root mean squared error* Tabel 9 yaitu 0.017623, sedangkan metode usulan Tabel 5 untuk fitur optimal ketiganya yaitu 0.016577, 0.016582, dan 0.016651. Hasil rata-rata *mean absolute error* Tabel 9 yaitu 0.0134993, sedangkan metode usulan Tabel 5 untuk fitur optimal ketiganya yaitu 0.012843, 0.012862, dan 0.012876. Hasil rata-rata *mean absolute percentage error* Tabel 9 yaitu 1.493767%, sedangkan metode usulan Tabel 5 untuk fitur optimal ketiganya yaitu 1.42236%, 1.42411%, dan 1.42616%.

Table 9 Peramalan menggunakan Multi-Layer Perceptron Regressor (Vijh et al., 2020)

Metrik	CV1	CV2	CV3	CV4	CV5	Means
MSE	0.0001659	0.00176	0.0005533	0.0011806	0.0021067	0.0011533
RMSE	0.012882	0.0419523	0.0235226	0.0343593	0.0458991	0.0317231
MAE	0.0099484	0.0379828	0.0181873	0.0288728	0.0427691	0.0275521
MAPE	1.068780%	4.188809%	2.143306%	3.118260%	4.816896%	3.067210%

Tabel 10 merupakan peramalan menggunakan *machine learning multi-layer perceptron regressor* tanpa seleksi fitur. Hasil rata-rata pengukuran *mse*, *rmse*, *mae*, dan *mape* yaitu 0.0011533, 0.0317231, 0.0275521, dan 3.067210%. Hasil pengukuran *mean squared error*, *root mean squared error*, *mean absolute error*, dan *mean absolute percentage error* Tabel 10 dibandingkan metode usulan pada Tabel 6 menunjukkan bahwa metode usulan lebih unggul pada semua pengukuran. Hasil rata-rata *mean squared error* pada Tabel 10 yaitu 0.0011533, sedangkan metode usulan pada Tabel 6 yaitu 0.000317, 0.00032, dan 0.000325. Hasil rata-rata *root mean squared error* Tabel 10 yaitu 0.0317231, sedangkan metode usulan pada Tabel 6 yaitu 0.017643, 0.017693, dan 0.01776. Hasil rata-rata *mean absolute error* Tabel 10 yaitu 0.0275521, sedangkan metode usulan pada Tabel 6 yaitu 0.013649, 0.013802, dan 0.013821. Hasil rata-rata *mean absolute percentage error* Tabel 10 yaitu 3.067210%, sedangkan metode usulan pada Tabel 6 yaitu 1.522516%, 1.527004%, dan 1.541950%.

Table 10 Peramalan menggunakan Random Forest Regression (Vijh et al., 2020)

Metrik	CV1	CV2	CV3	CV4	CV5	Means
MSE	0.0004266	0.0012858	0.0007337	0.0003654	0.0003101	0.0006243
RMSE	0.0206537	0.0358582	0.027086	0.0191149	0.017609	0.0240644
MAE	0.015605	0.0323968	0.0215943	0.0157527	0.0139964	0.019869
MAPE	1.655635%	3.581127%	2.545511%	1.724639%	1.570199%	2.215423%

Tabel 11 merupakan peramalan menggunakan *machine learning random forest regression* tanpa seleksi fitur. Hasil rata-rata *mean squared error*, *root mean squared error*, *mean absolute error*, dan *mean absolute percentage error* yaitu 0.0006243, 0.0240644, 0.019869, dan 2.215423%. Hasil pengukuran pada Tabel 11 dibandingkan dengan metode usulan pada Tabel 7 menunjukkan bahwa metode usulan lebih unggul dari seluruh pengukuran yang dilakukan. Hasil rata-rata *mean squared error* Tabel 11 yaitu 0.0006243, sedangkan metode usulan pada Tabel 7 yaitu 0.000459, 0.000489, dan 0.000553. Hasil rata-rata *root mean squared error* Tabel 11 yaitu 0.0240644, sedangkan metode usulan pada Tabel 7 yaitu 0.021248, 0.021788, dan 0.02286. Hasil rata-rata *mean absolute error* Tabel 11 yaitu 0.019869, sedangkan metode usulan pada Tabel 7 yaitu 0.016906, 0.017309, dan 0.018782. Hasil rata-rata *mean absolute percentage error* table 11 yaitu 2.215423%, sedangkan metode usulan Tabel 7 yaitu 1.867293%, 1.909080%, dan 2.095415%.

Table 11 Peramalan menggunakan Support Vector Regression (Xia et al., 2013)

Metrik	CV1	CV2	CV3	CV4	CV5	Means
MSE	0.01091764	0.0040696	0.0005899	0.0013354	0.000447	0.0034719
RMSE	0.10448753	0.0637935	0.024287	0.0365437	0.0211412	0.0500506
MAE	0.10275549	0.0581385	0.0202751	0.0323792	0.0176658	0.0462428
MAPE	10.981925%	6.306595%	2.386893%	3.492601%	1.966023%	5.026808%

Tabel 12 merupakan peramalan menggunakan *machine learning support vector regression* tanpa seleksi fitur. Hasil rata-rata *mean squared error*, *root mean squared error*, *mean absolute error*, dan *mean absolute percentage error* yaitu 0.0034719, 0.0500506, 0.0462428, dan 5.026808%. Hasil pengukuran pada Tabel 12 dibandingkan dengan metode usulan pada Tabel 8 menunjukkan bahwa metode usulan lebih unggul pada semua pengukuran yang dilakukan. Hasil rata-rata *mean squared error* Tabel 12 yaitu 0.0034719, sedangkan metode usulan pada Tabel 8 yaitu 0.003044, 0.003083, dan 0.003084. Hasil rata-rata pengukuran *root mean squared error* Tabel 12 yaitu 0.0500506, sedangkan metode usulan pada Tabel 8 yaitu 0.045788, 0.045955, dan 0.046067. Hasil rata-rata pengukuran *mean absolute error* Tabel 12 yaitu 0.0462428, sedangkan hasil metode usulan pada Tabel 8 yaitu 0.042367, 0.042407, dan 0.042544. Hasil rata-rata pengukuran *mean absolute percentage error* Tabel 12 yaitu 5.026808%, sedangkan hasil metode usulan pada Tabel 8 yaitu 4.594368%, 4.596041%, dan 4.618072%.

4. CONCLUSION

Penelitian ini mengusulkan fitur penting berbasis impurity untuk seleksi fitur pada recursive feature elimination pada Peramalan Harga Saham dengan data historical dan various moving average. Data saham yang digunakan data transaksi harian dan data tekniska indikator various moving average (simple moving average, exponenetial moving average, dan real moving average). Seleksi fitur menggunakan fitur penting berbasis impurity pada recursive feature elimination. Peramalan menggunakan machine learning linear regression, multi-layer perceptron regressor, random forest regressor, dan support vector regression. Hasil pengukuran mse, rmse, mae, dan mape fitur optimal paling baik didapatkan dengan jumlah 6 fitur dan machine learning linear regression. Nilai rata-rata mse, rmse, mae, dan mape yaitu 0.000279, 0.016577, 0.012843, dan 1.42236%. Hal ini membuktikan fitur penting berbasis impurity untuk seleksi fitur pada recursive feature elimination dengan menggunakan data histori dan various moving average berhasil melakukan peramalan.

5. ACKNOWLEDGMENTS

Terimakasih kepada Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi atas dukungan finansial melalui skema Penelitian Dosen Pemula (PDP) tahun 2023 dengan nomor kontrak 074 / E5 / PG.02.00.PM / 2023, 024 / LL10 / PG.AK / 2023, 119 / LPPM / KPD / DPRM / VIII / 2023. Terimakasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat Universitas Pahlawan Tuanku Tambusai yang telah memberikan arahan bagi dosen pemula dalam melaksanakan skema penelitian dosen pemula.

6. REFERENCES

- Albahli, S., Nazir, T., Mehmood, A., Irtaza, A., Alkhalifah, A., & Albattah, W. (2022). AEI-DNET: A Novel DenseNet Model with an Autoencoder for the Stock Market Predictions Using Stock Technical Indicators. *Electronics (Switzerland)*, 11(4), 611. <https://doi.org/10.3390/electronics11040611>
- Ampomah, E. K., Nyame, G., Qin, Z., Addo, P. C., Gyamfi, E. O., & Gyan, M. (2021). Stock Market Prediction with Gaussian Naïve Bayes Machine Learning Algorithm. *Informatica*, 45(2), 243–256. <https://doi.org/10.31449/inf.v45i2.3407>

- Ayyappa, Y., & Siva Kumar, A. P. (2023). Optimized long short-term memory-based stock price prediction with sentiment score. *Social Network Analysis and Mining*, 13(1), 13. <https://doi.org/10.1007/s13278-022-01004-5>
- Azimifar, M., Araabi, B. N., & Moradi, H. (2020). Forecasting stock market trends using support vector regression and perceptually important points. *2020 10th International Conference on Computer and Knowledge Engineering (ICCCKE)*, 268–273. <https://doi.org/10.1109/ICCCKE50421.2020.9303667>
- Azizah, M., Irawan, M. I., & Putri, E. R. M. (2020). Comparison of stock price prediction using geometric Brownian motion and multilayer perceptron. *AIP Conference Proceedings*, 2242(May), 030016. <https://doi.org/10.1063/5.0008066>
- Basak, S., Kar, S., Saha, S., Khaidem, L., & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, 47, 552–567. <https://doi.org/10.1016/j.najef.2018.06.013>
- Bhuriya, D., Kaushal, G., Sharma, A., & Singh, U. (2017). Stock market predication using a linear regression. *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, 510–513. <https://doi.org/10.1109/ICECA.2017.8212716>
- Bouktif, S., Fiaz, A., & Awad, M. (2019). Stock Market Movement Prediction using Disparate Text Features with Machine Learning. *2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS)*, 1–6. <https://doi.org/10.1109/ICDS47004.2019.8942303>
- Chandrika, P. V., & Srinivasan, K. S. (2021). Predicting Stock Market Movements Using Artificial Neural Networks. *Universal Journal of Accounting and Finance*, 9(3), 405–410. <https://doi.org/10.13189/ujaf.2021.090315>
- Charan, V. S., Rasool, A., & Dubey, A. (2022). Stock Closing Price Forecasting using Machine Learning Models. *2022 International Conference for Advancement in Technology, ICONAT 2022*. <https://doi.org/10.1109/ICONAT53423.2022.9725964>
- Chauhan, K., & Sharma, N. (2023). Study of Linear Regression Prediction Model for American Stock Market Prediction. In *Advances in Transdisciplinary Engineering* (Vol. 32, pp. 406–411). <https://doi.org/10.3233/ATDE221289>
- Deshmukh, R. A., Jadhav, P., Shelar, S., Nikam, U., Patil, D., & Jawale, R. (2023). Stock Price Prediction Using Principal Component Analysis and Linear Regression. *Lecture Notes in Networks and Systems*, 490, 269–276. https://doi.org/10.1007/978-981-19-4052-1_28
- Fang, Y. (2020). Stock Price Forecasting Based on Improved Support Vector Regression. *Proceedings - 2020 7th International Conference on Information Science and Control Engineering, ICISCE 2020*, 1351–1354. <https://doi.org/10.1109/ICISCE50968.2020.00272>
- Ghosh, P., Neufeld, A., & Sahoo, J. K. (2022). Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. *Finance Research Letters*, 46, 102280. <https://doi.org/10.1016/j.frl.2021.102280>
- Hamzah, S. R., Halul, H., Jeng, A., & Umul Ain'syah Sha'ari. (2021). Forecasting Nestle Stock Price by using Brownian Motion Model during Pandemic Covid-19. *Malaysian Journal of Science Health & Technology*, 7(2), 58–64. <https://doi.org/10.33102/mjosht.v7i2.214>
- Huang, Y., Deng, C., Zhang, X., & Bao, Y. (2022). Forecasting of stock price index using support vector regression with multivariate empirical mode decomposition. *Journal of Systems and Information Technology*, 24(2), 75–95. <https://doi.org/10.1108/JSIT-12-2019-0262>
- Illa, P. K., Parvathala, B., & Sharma, A. K. (2022). Stock price prediction methodology using random forest algorithm and support vector machine. *Materials Today: Proceedings*, 56, 1776–1782. <https://doi.org/10.1016/j.matpr.2021.10.460>
- Ji, Y., Liew, A. W.-C., & Yang, L. (2021). A Novel Improved Particle Swarm Optimization With Long-Short Term Memory Hybrid Model for Stock Indices Forecast. *IEEE Access*, 9, 23660–23671. <https://doi.org/10.1109/ACCESS.2021.3056713>
- K, P., Rudagi, S., M, N., Patil, R., & Wadi, R. (2021). Comparative Study: Stock Prediction Using Fundamental and Technical Analysis. *2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNBC)*, 1–4. <https://doi.org/10.1109/ICMNBC52512.2021.9688449>
- Kehinde, T. O. O., Chan, F. T. S., & Chung, S. H. H. (2023). Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey. *Expert Systems with Applications*, 213, 119299. <https://doi.org/10.1016/j.eswa.2022.119299>
- Kumar, G., Jain, S., & Singh, U. P. (2021). Stock Market Forecasting Using Computational Intelligence: A Survey. *Archives of Computational Methods in Engineering*, 28(3), 1069–1101. <https://doi.org/10.1007/s11831-020-09413-5>

- Lavingia, K., Khanpara, P., Mehta, R., Patel, K., & Kothari, N. (2022). Predicting Stock Market Trends using Random Forest: A Comparative Analysis. *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 1544–1550. <https://doi.org/10.1109/ICCES54183.2022.9835876>
- Lee, M.-C., Chang, J.-W., Yeh, S.-C., Chia, T.-L., Liao, J.-S., & Chen, X.-M. (2022). Applying attention-based BiLSTM and technical indicators in the design and performance analysis of stock trading strategies. *Neural Computing and Applications*, *34*(16), 13267–13279. <https://doi.org/10.1007/s00521-021-06828-4>
- Ma, Y., Mao, R., Lin, Q., Wu, P., & Cambria, E. (2023). Multi-source aggregated classification for stock price movement prediction. *Information Fusion*, *91*, 515–528. <https://doi.org/10.1016/j.inffus.2022.10.025>
- Meghana, N., & Arumugam, S. S. (2023). Use Novel Naive Bayes to Predict Stock Market Movement and Compare Prediction Accuracy to Linear Regression. *Journal of Survey in Fisheries Sciences*, *10*(1S), 2874–2882. <http://sifisheriesciences.com/journal/index.php/journal/article/view/520>
- Namdari, A., & Durrani, T. S. (2021). A Multilayer Feedforward Perceptron Model in Neural Networks for Predicting Stock Market Short-term Trends. *Operations Research Forum*, *2*(3), 38. <https://doi.org/10.1007/s43069-021-00071-2>
- Pandya, J. B., & Jaliya, U. K. (2022). An empirical study on the various stock market prediction methods. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, *8*(1), 58–80. <https://doi.org/10.26594/register.v8i1.2533>
- Park, H. J., Kim, Y., & Kim, H. Y. (2022). Stock market forecasting using a multi-task approach integrating long short-term memory and the random forest framework. *Applied Soft Computing*, *114*, 108106. <https://doi.org/10.1016/j.asoc.2021.108106>
- Ray, R., Khandelwal, P., & Baranidharan, B. (2018). A Survey on Stock Market Prediction using Artificial Intelligence Techniques. *2018 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 594–598. <https://doi.org/10.1109/ICSSIT.2018.8748680>
- Sundar, G., & Satyanarayana, K. (2019). Multi Layer Feed Forward Neural Network Knowledge Base to Future Stock Market Prediction. *International Journal of Innovative Technology and Exploring Engineering*, *8*(11S), 1061–1075. <https://doi.org/10.35940/ijitee.K1218.09811S19>
- Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock Closing Price Prediction using Machine Learning Techniques. *Procedia Computer Science*, *167*(2019), 599–606. <https://doi.org/10.1016/j.procs.2020.03.326>
- Xia, Y., Liu, Y., & Chen, Z. (2013). Support Vector Regression for prediction of stock trend. *2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering*, 123–126. <https://doi.org/10.1109/ICIM.2013.6703098>
- Yeo, L. L. X., Cao, Q., & Quek, C. (2023). Dynamic portfolio rebalancing with lag-optimised trading indicators using SeroFAM and genetic algorithms. *Expert Systems with Applications*, *216*, 119440. <https://doi.org/10.1016/j.eswa.2022.119440>