

Komparasi Model Gerak *Brown* Geometrik Termodifikasi dan Model Kecerdasan Buatan untuk Prediksi Harga Saham Sektor Kesehatan di Indonesia

RR. Kurnia Novita Sari^{1*}, Wilan Sutisna², Martha Joanadiva Majesty Wororomi³, Venansius Ryan Tjahjono⁴ 

¹ Kelompok Keahlian Statistika, Institut Teknologi Bandung, Bandung, Indonesia

^{2,3,4} Magister Aktuaria, Institut Teknologi Bandung, Bandung, Indonesia

ARTICLE INFO

Article history:

Received June 27, 2022

Revised June 30, 2022

Accepted February 22, 2023

Available online July 25, 2023

Kata Kunci:

harga Saham, Gerak *Brown* Geometrik, Kecerdasan Buatan

Keywords:

Stock Prices, Geometric Brownian Motion, Artificial Intelligence



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2023 by Author. Published by Universitas Pendidikan Ganesha.

ABSTRAK

Selama pandemi COVID-19 di Indonesia, pergerakan saham di berbagai sektor mengalami dampak yang signifikan. Akan tetapi, harga saham pada sektor kesehatan mengalami pertumbuhan sebesar 3,71%. Penelitian ini bertujuan menerapkan model Gerak *Brown* Geometrik yang Termodifikasi (MBM) dan model kecerdasan buatan untuk memprediksi pergerakan penutupan harga pada data harga penutupan saham PT. Kalbe Farma Tbk. Lalu, hasil prediksi dibandingkan untuk mengetahui model terbaik. Pada MBM, dilakukan konstruksi tren dan volatilitas yang memuat variabel waktu. Hasil ini kemudian dibandingkan dengan model kecerdasan buatan, yaitu Light Gradient Boosting Machine (LightGBM) dan Multi Layer Perceptron (MLP). Selain itu, pada konstruksi model kecerdasan buatan, dilakukan pendekatan baru yang mana input dari model tersebut adalah tanggal dan titik urutan deret waktu. Hasil simulasi menunjukkan bahwa MBM mampu memberikan nilai MAPE yang lebih kecil dibandingkan GBM. Akan tetapi, model kecerdasan buatan, khususnya LightGBM, masih lebih unggul bila digunakan untuk memodelkan harga saham sektor kesehatan. Model ini memberikan nilai MAPE dibawah 1%.

ABSTRACT

During the COVID-19 pandemic in Indonesia, the stock price movement experienced a significant loss. Nevertheless, the stock prices in the health sector show an increase of 3.71%. This study intends to apply the Modified Geometric Brownian Motion (MBM) and machine learning model to build a predictive model on PT. Kalbe Farma Tbk. data. The predicted results are then compared to give an insight into the best model for predicting the stock price in the health sector in Indonesia. The proposed novelty for MBM are constructing the trend and volatility that depends on the time. This model is then resembled with the artificial intelligence models, which are Light Gradient Boosting Machine (LightGBM) and Multi Layer Perceptron (MLP). Furthermore, we also apply a novel modification for the AI-based model. We propose the input of the artificial intelligence model is the date and time series sequence point. The simulation results show the MBM is able to provide a smaller MAPE value than the GBM. Moreover, an artificial intelligence model, specifically LightGBM, shows a superiority when applied to model the stock prices. This model presents a MAPE value below 1%.

1. PENDAHULUAN

Pandemi COVID-19 terjadi di Indonesia sejak awal 2020 hingga saat ini telah berdampak secara signifikan pada hampir seluruh sektor ekonomi. Dalam jangka pendek, banyak negara menerapkan kebijakan karantina ketat yang mengakibatkan kegiatan ekonomi menjadi sangat terbatas dan jangka panjangnya dapat meningkatkan jumlah pengangguran dan kegagalan bisnis seperti pada sektor pariwisata dan penerbangan (D, M., & Q., 2020; Ichev, 2021). Berbeda dengan sektor kesehatan, Badan Pusat Statistik (BPS) mencatat Produk Domestik Bruto (PDB) di sektor kesehatan mengalami pertumbuhan secara *y-o-y* (*year of year*) di setiap triwulan. Pada triwulan kedua 2020, sektor kesehatan mengalami pertumbuhan sebesar 3,71% terhadap triwulan kedua 2019 (*y-o-y*). Hal ini berbanding terbalik dengan pertumbuhan ekonomi Indonesia di sektor lain yang mengalami kontraksi pertumbuhan sebesar 5,32%. Tumbuhnya PDB sektor kesehatan beriringan dengan minat investor untuk menanamkan modal pada perusahaan kesehatan. Hal ini sejalan dengan pergerakan indeks saham *IDX Healthcare*. Sejak dirilis pada 25 Januari 2021, *IDX Healthcare* terdapat 21 emiten, antara lain: Kimia Farma Tbk., Kalbe Farma Tbk., Industri Jamu dan Farmasi

*Corresponding author.

E-mail addresses: kurnia@math.itb.ac.id (RR. Kurnia Novita Sari)

Sido Muncul Tbk. Indeks saham gabungan sektor kesehatan ini mencapai pertumbuhan sebesar 17,8% y-o-y di 2020.

Investor melihat kontribusi sektor kesehatan dapat memberikan keuntungan investasi di masa pandemi COVID-19. Pada awal 2021, jumlah pasien terinfeksi COVID-19 di Indonesia mengalami peningkatan. PT. Kalbe Farma Tbk. melihat bahwa penjualan vitamin dan suplemen tumbuh antara 10% dan 20% setiap tahun sehingga mereka memanfaatkan *platform* digital seperti klikdokter. Selain itu, PT. Kalbe Farma Tbk. juga bekerja sama dengan perusahaan Bioteknologi Korea Selatan dalam mengembangkan vaksin menggunakan *platform* virus DNA dan pengujian klinis kandidat vaksin COVID-19 di Indonesia. PT. Kalbe Farma Tbk. mencatat kenaikan pendapatan pada triwulan pertama 2021, hal ini didukung oleh peningkatan penjualan di segmen jasa distribusi, obat resep, dan produk kesehatan. Pandemi COVID-19 ini telah membawa ketidakpastian terutama pada pergerakan harga saham dengan tingkat volatilitas yang tinggi (Ashraf, 2021; Zhang, Ding, Hang, & He, 2022). Oleh karena itu, prediksi saham ini menjadi penting sehingga perkiraan harga saham di masa yang akan datang dapat diketahui dan tidak membawa kerugian besar bagi investor. Prediksi saham bertujuan untuk memberikan gambaran kepada investor dan pemegang saham mengenai pergerakan aset. Namun, prediksi saham terkadang cukup sulit disebabkan adanya faktor eksternal yang dapat menimbulkan fluktuasi harga saham sehingga volatilitasnya tinggi (Ruslan & Mokhtar, 2021; Salisu & Vo, 2020).

Pergerakan saham dapat dimodelkan secara matematis menggunakan model probabilistik dan juga kecerdasan buatan (Na & Kim, 2021; Syaifudin & Putri, 2021). Pada penelitian ini, dibuat beberapa jenis model prediksi yang berbasis model konvensional runtun waktu dan model yang berbasis kecerdasan buatan. Model runtun waktu yang bisa digunakan adalah ARIMA dan GARCH. Model ini memiliki kemampuan untuk menangkap fenomena tren (kecenderungan naik atau turun) dan volatilitas. Akan tetapi, kedua model hanya bisa digunakan untuk memprediksi satu langkah saja (*one-step forecast*). Dengan demikian, pengamatan secara jangka panjang (*long term*) tidak bisa dilakukan dengan model ini. Sehingga pada penelitian ini diusulkan model stokastik Gerak Brown Geometrik (GBM) yang memuat faktor tren dan volatilitas (Feng, Jiang, & Volkmer, 2021; Hull, 2018). Faktor tren dan volatilitas pada GBM dibuat konstan yang menjadi kelemahan pada model ini. Terdapat beberapa penelitian untuk mengantisipasi kelemahan model GBM tersebut. Salah satunya adalah model hibrid GBM dengan metode *Predictive Control*. Model hibrid ini mampu meningkatkan akurasi dari GBM tetapi belum cukup baik dan menimbulkan fenomena *out sample* (Syaifudin & Putri, 2021). Oleh sebab itu, dikonstruksi model Gerak Brown Geometrik Termodifikasi (MBM) yang mengubah faktor tren dan volatilitas konstan menjadi suatu fungsi terhadap waktu. Fungsi tren memuat beberapa jenis fungsi, sedangkan fungsi volatilitas menggunakan model 3/4-volatilitas. Model 3/4-volatilitas mampu memberikan performansi yang lebih baik dibandingkan dengan faktor volatilitas pada GBM yang kemudian digunakan untuk valuasi opsi keuangan (Oud & Goard, 2015; Tong & Liu, 2018).

Hasil pemodelan MBM dianalisis dan dikomparasikan dengan model berbasis kecerdasan buatan. Beberapa model berbasis kecerdasan buatan yang bisa digunakan untuk memodelkan pergerakan aset adalah model jaringan saraf tiruan (Na & Kim, 2021; Yu, Qin, Chen, & Parmar, 2020; Zolfaghari & Gholami, 2021) dan model pohon keputusan (Sun, Liu, & Sima, 2020; Zhou, Zhang, Sornette, & Jiang, 2019). Pada penelitian yang sudah ada, data yang digunakan sebagai *input* adalah harga buka, harga tertinggi, harga terendah, dan volume pasar aset (Daul, Jaisson, & Nagy, 2022; Kamara, Chen, E., & Pan, 2022; Tao, Gao, Mao, & Huang, 2022). Namun, pada saat digunakan untuk melakukan prediksi jangka pendek, data harga buka, harga tinggi, harga rendah, dan volume pasar aset tidak diketahui. Untuk melakukan prediksi jangka panjang, dibuat suatu mekanisme baru yaitu model kecerdasan buatan berbasis komponen tanggal (tanggal, hari, minggu, bulan, dan deret). Pembentukan data *input* melibatkan skema *one-hot encoding*. Kemudian, model berbasis kecerdasan yang digunakan adalah *Light Gradient Boosting Machine* (LightGBM) dan *Multi Layer Perceptron* (MLP). Model LightGBM merupakan salah satu jenis pohon keputusan *gradient boosting* yang bisa digunakan di berbagai jenis data dan berpotensi melampaui model kecerdasan buatan lainnya dalam hal prediksi harga (Jabeur, Khalfaoui, & Arfi, 2021; Qu, Zhang, & Qin, 2020; Sun et al., 2020). Sedangkan, model MLP dipilih karena merupakan model komputasi non-linier yang mampu mempelajari perilaku data dengan hasil yang memuaskan (Chhajer, Shah, & Kshirsagar, 2022; Huang et al., 2021; Khang et al., 2021; Rezaei, Faaljoui, & Mansourfar, 2021).

Pada artikel ilmiah ini, latar belakang mengenai alasan dan tujuan penelitian dijelaskan pada Bagian 1. Bagian 2 memuat metode yang digunakan yaitu Gerak Brown Geometrik, Gerak Brown Geometrik Termodifikasi, *Light Gradient Boosting Machine*, dan *Multi Layer Perceptron*. Keempat model tersebut akan diaplikasikan pada data saham PT. Kalbe Farma Tbk. Pada Bagian 3 akan diuraikan deskripsi data, visualisasi, hasil, dan pembahasan. Kesimpulan dijelaskan pada Bagian 4.

2. METODE

Pada penelitian ini, data pergerakan harga saham kesehatan yang akan digunakan adalah harga penutupan saham PT. Kalbe Farma Tbk. Data yang diambil sebanyak 489 data yaitu pada periode Januari 2020 s.d Desember 2021. Adapun metode analisis yang digunakan untuk memprediksi harga penutupan saham, yaitu model matematis konvensional (Gerak Brown Geometrik dan Gerak Brown Geometrik Termodifikasi) dan model kecerdasan buatan (*Light Gradient Boosting Machine* dan *Multi Layer Perceptron*). Kedua jenis model tersebut akan dianalisis dan dibandingkan, kemudian dipilih model terbaik untuk memprediksi harga saham. Simulasi model dilakukan menggunakan perangkat lunak Python v3.7.6 dan MATLAB R2021a 64bit, *Student License*.

Gerak Brown Geometrik

Harga saham dapat dipandang sebagai barisan peubah acak $\{S_t, t \geq 0\}$ dengan indeks parameter waktu t . Harga saham ini merupakan salah satu realisasi dari data runtun waktu stokastik pada sektor keuangan. Salah satu model matematis yang bisa digunakan untuk melakukan valuasi dan prediksi harga saham adalah Gerak Brown Geometrik (*Geometric Brownian Motion*), yang disebut GBM. GBM untuk proses pergerakan harga saham diformulasikan sebagai suatu persamaan diferensial stokastik yaitu (Hull, 2018).

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{1}$$

dengan dS_t adalah perubahan harga saham S_t pada interval waktu dt dan W_t adalah Gerak Brown standar yang berdistribusi normal standar. Parameter μ dan σ masing-masing menyatakan tren dan volatilitas dari harga saham. Dengan menerapkan Lema Ito pada Persamaan 1, diperoleh

$$S_t = S_0 \exp \left(\left[\mu - \frac{\sigma^2}{2} \right] t + \sigma W_t \right). \tag{2}$$

Persamaan 2 adalah solusi dari persamaan diferensial stokastik GBM yang memberikan solusi analitik dari pergerakan acak harga saham. Tinjau, bahwa pada Persamaan 2, terdapat faktor Gerak Brown yang memiliki distribusi normal standar. Hal ini berarti,

$$\ln \left(\frac{S_t}{S_0} \right) \sim N \left(\left[\mu - \frac{\sigma^2}{2} \right] t, \sigma^2 t \right). \tag{3}$$

Untuk mengestimasi parameter μ dan σ pada Persamaan 3, digunakan metode *Maximum Likelihood Estimation* (MLE) yang estimatornya sebagai berikut (Hull, 2018)

$$\hat{\mu} = \frac{\sum_{t=1}^n r_t}{n \Delta t} + \frac{\hat{\sigma}^2}{2\Delta t} \quad \text{dan} \quad \hat{\sigma}^2 = \sqrt{\frac{1}{(n-1) \Delta t} \sum_{t=1}^n (r_t - \bar{r})^2}. \tag{4}$$

Gerak Brown Geometrik Termodifikasi

Kelemahan dari model GBM terletak pada μ dan σ yang merupakan konstanta. Konstanta ini membatasi pergerakan tren dan volatilitas (Putri, Tjahjono, & Imron, 2021). Model GBM ini dapat diperumum dengan membuat μ dan σ menjadi suatu fungsi mean dan volatilitas sehingga

$$dS_t = \theta(\mu, S_t) dt + \xi(\sigma, S_t) dW_t \tag{5}$$

dengan $\theta(\mu, S_t)$ dan $\xi(\sigma, S_t)$ adalah fungsi mean dan volatilitas yang memuat parameter GBM dan harga saham. Aba Oud dan Goard memodifikasi faktor tren dan volatilitas menjadi (Oud & Goard, 2015).

$$\theta(\mu, S_t) = a\sqrt{S_t} + bS_t \tag{6}$$

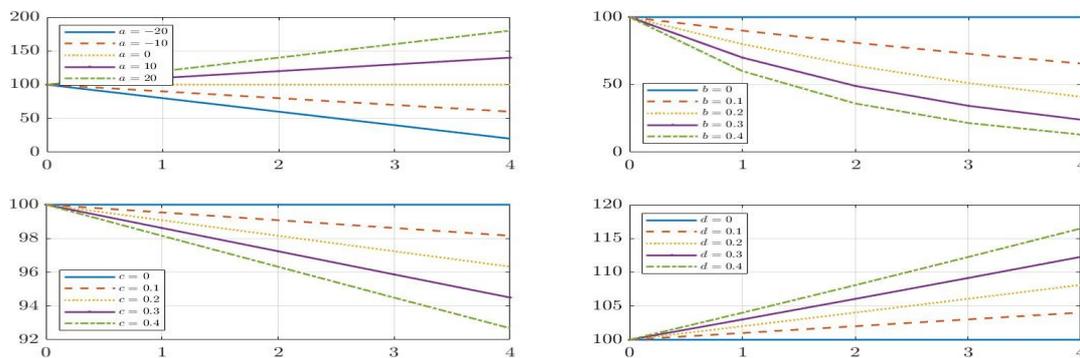
$$\xi(\sigma, S_t) = \sigma S_t^{\frac{3}{4}}. \tag{7}$$

Hasil prediksi yang melibatkan model 3/4-volatilitas memberikan hasil yang lebih akurat dibandingkan dengan model stokastik yang umum digunakan (Oud & Goard, 2015; Tong & Liu, 2018). Pada

penelitian ini, dilakukan perumuman pada fungsi tren dengan menambahkan dua jenis faktor baru yaitu suku konstan dan suku logaritma natural menjadi

$$\theta(\mu, S_t) = a + bS_t + c \ln(S_t) + d\sqrt{S_t}. \tag{8}$$

Suku konstan akan memberikan suatu kapabilitas untuk mendeteksi adanya perubahan tren secara linier. Sedangkan faktor logaritma natural akan menambahkan efek perubahan imbal hasil terhadap tren perubahan harga saham secara logaritmik. Adapun diberikan perilaku geometri dari masing-masing parameter pada tren yang dimodifikasi yang diberikan pada Gambar 1. Parameter a memberikan sifat linier-simetris pada pergerakan harga saham. Parameter b menambahkan unsur 'kelengkungan' pada pergerakan tren saham. Sedangkan parameter c dan d masing-masing bersifat kebalikan dengan parameter c mengontrol pergerakan tren ke arah bawah dan d mengontrol pergerakan tren ke arah atas. Dengan demikian, keempat parameter yang ditambahkan mampu memberikan sifat dinamis pada pergerakan tren saham. Sifat ini diperlukan agar pergerakan prediksi saham bisa semakin akurat dan mendekati harga aktualnya.



Gambar 1. Perubahan Parameter a, b, c, d terhadap Pergerakan Tren Harga Saham

Untuk faktor volatilitas, kami mengadopsi model 3/4-volatilitas sebab model ini mampu menjelaskan suatu proses *mean-reverting* yang menggambarkan penurunan harga secara cepat setelah adanya lonjakan pada harga (Oud & Goard, 2015). Dengan demikian, model GBM Termodifikasi yang digunakan adalah

$$d\tilde{S}_t = \left(a + b\tilde{S}_t + c \ln(\tilde{S}_t) + d\sqrt{\tilde{S}_t} \right) dt + \sigma\tilde{S}_t^{\frac{3}{4}} d\tilde{W}_t \tag{9}$$

dengan \tilde{W}_t merupakan Gerak Brown standar. Untuk menyelesaikan persamaan diferensial stokastik pada Persamaan 9, diterapkan Lema Ito yang diberikan oleh

$$dP_t = \left[g(\tilde{S}_t, t) \frac{\partial P_t}{\partial \tilde{S}_t} + \frac{\partial P_t}{\partial t} + \frac{h(\tilde{S}_t, t)}{2} \frac{\partial^2 P_t}{\partial \tilde{S}_t^2} \right] dt + \left[h(\tilde{S}_t, t) \frac{\partial P_t}{\partial \tilde{S}_t} \right] d\tilde{W}_t. \tag{10}$$

Namun, bentuk persamaan yang diperoleh tidaklah *closed-form* sehingga untuk mendapatkan bagaimana pergerakan \tilde{S}_t harus menggunakan simulasi numerik. Kemudian, untuk melakukan estimasi parameter pada MBM, digunakan teknik *Hyperparameter Optimization* yaitu algoritma *Adaptive Random Search* (ARS). Algoritma ini akan mencari secara acak kombinasi parameter terbaik yang meminimumkan fungsi *fitness* berikut

$$f(\mathbf{p}) = \sqrt{\frac{1}{n} \sum_{t=1}^n (\tilde{S}_t^{pred}(\mathbf{p}) - \tilde{S}_t^{akt})^2} \tag{11}$$

dengan n banyaknya titik observasi, $\tilde{S}_t^{pred}(\mathbf{p})$ merupakan nilai prediksi yang diberikan oleh Persamaan 9 dengan vektor parameter \mathbf{p} , dan \tilde{S}_t^{akt} adalah data aktual. Algoritma untuk mendapatkan

melakukan estimasi parameter dari MBM yang diadopsi dari penelitian disajikan pada [Tabel 1](#). (Konstantinov, Diveev, Balandina, & Baryshnikov, 2019).

Tabel 1. Algoritma *Adaptive Random Search*

Input	<ol style="list-style-type: none"> 1. Tetapkan M sebagai banyak iterasi. 2. Tetapkan m sebagai banyak populasi awal. 3. Bangkitkan vektor $a \sim U_m(0,85S_0; 1,15S_0)$, $b \sim U_m(0; 0,0002)$, $c \sim U_m(0; 0,0002)$, $d \sim U_m(0; 0,0002)$, dan $\sigma \sim U_m(0; 0,125)$. 4. Tetapkan f_0 sebagai nilai awal fungsi <i>fitness</i>, $f(\mathbf{p})$.
Proses	<ol style="list-style-type: none"> 1. Pilih $\mathbf{p}^{(i)} = (a^{(i)}, b^{(i)}, c^{(i)}, d^{(i)}, \sigma^{(i)})$ secara acak dari data bangkitan. 2. Hitung nilai $f(\mathbf{p}^{(i)})$. 3. Apabila $f(\mathbf{p}^{(i)}) < h_0$, simpan nilai $\mathbf{p}^{(i)}$ dan perbarui $f_0 = f(\mathbf{p}^{(i)})$. 4. Ulangi langkah proses 1 sampai dengan 3 sebanyak M kali.
Output	$\arg \min f(\mathbf{p}^{(i)})$

Light Gradient Boosting Machine

Machine Learning (ML) merupakan pendekatan dalam *Artificial Intelligence* (AI) yang banyak digunakan untuk menirukan perilaku manusia untuk menyelesaikan masalah atau melakukan otomatisasi. ML mencoba menirukan bagaimana proses manusia belajar dan menggeneralisasi. Adapun tahap prediksi digunakan oleh mesin untuk menerka *output* dari suatu data *input* berdasarkan data yang sudah dipelajari dalam *training* (Ke et al., 2017). Salah satu ML yang akan digunakan adalah *Light Gradient Boosting Machine* (LightGBM). LightGBM adalah model kerangka kerja peningkatan baru yang merupakan pengembangan dari *Gradient Boosting Decision Tree* (GBDT) yang merupakan algoritma pembelajaran mesin yang banyak digunakan karena efisiensi dan akurasi. Dengan munculnya *big data*, GBDT menghadapi tantangan baru, terutama dalam *tradeoff* antara efisiensi dan akurasi. Implementasi konvensional GBDT diperlukan, untuk setiap fitur, memindai semua contoh data untuk memperkirakan perolehan informasi dari semua kemungkinan *split points*. Hal ini menyebabkan implementasi ini membutuhkan waktu yang sangat lama saat menangani *big data* (Ke et al., 2017). Saat proses *training*, GBDT akan membuat serangkaian pengklasifikasian dasar (*decision tree*) dan menghitung parameter bobot dari setiap *classifier* dalam iterasi. Semua pengklasifikasian dasar dan bobotnya digabungkan untuk membentuk model akhir dan didefinisikan sebagai rumus

$$F_m(X) = \omega_1 f_1(X) + \omega_2 f_2(X) + \dots + \omega_m f_m(X) \tag{12}$$

dengan $f_m(X)$ adalah *base classifier* dan ω_m adalah bobot parameter dari setiap *classifier*. Dua teknik baru diperkenalkan: *Gradient-based One-Side Sampling* (GOSS) dan *Exclusive Feature Bundling* (EFB). GOSS merupakan metode yang menyimpan semua atribut data dengan gradien besar dan melakukan *random sampling* pada atribut data dengan gradien kecil. Pada saat menghitung perolehan informasi, GOSS meinput sebuah konstanta *multiplier* pada data dengan gradien kecil dengan tujuan untuk mengimbangi pengaruh dari distribusi data. EFB merupakan teknik yang bertujuan untuk mengelompokkan fitur-fitur eksklusif untuk mengurangi jumlah fitur tanpa merusak akurasi penentuan *split points* (Ke et al., 2017). LightGBM merupakan salah satu ML yang sering digunakan untuk memprediksi harga saham dan terbukti kinerjanya bagus (Jabeur et al., 2021) dengan nilai *error* yang sangat kecil (Sun et al., 2020). Algoritma Light GBM dapat dilihat pada [Tabel 2](#).

Tabel 2. Algoritma LightGBM

Input	<ol style="list-style-type: none"> 1. Tetapkan M sebagai banyak iterasi. 2. <i>Training</i> data $X = (\vec{x}_i, y_i)_{i=1}^n$, dengan $\vec{x}_i = (x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, x_i^{(4)}, x_i^{(5)})$
Proses	<ol style="list-style-type: none"> 1. Hitung nilai $f_T(X) = \sum_{t=1}^T f_t(X)$ 2. Hitung nilai $\Gamma_t = \sum_{i=1}^n L(y_i, F_{t-1}(x_i) + f_t(x_i))$. 3. Perbaharui bobot $\omega_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i \lambda}$. 4. Perbaharui $\Gamma_t^* = -\frac{1}{2} \sum_{j=1}^J \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda}$. 5. Ulangi langkah proses 1 sampai dengan 4 sebanyak M kali.
Output	$\arg \min(\Gamma_t^*)$

Multi Layer Perceptron

Deep Learning (DL) adalah jaringan saraf tiruan yang menggunakan teknik tertentu untuk mempercepat proses pembelajaran dengan menggunakan beberapa *layer* atau lebih. *Multi Layer Perceptron* (MLP) adalah salah satu jaringan saraf *feed forward neural network* yang sederhana dan paling efisien yang dapat digunakan untuk melakukan prediksi, meminimalkan kesalahan di setiap iterasi hingga nilai yang diinginkan tercapai (Tiwari, Bharadwaj, & Gupta, 2017). MLP yang paling sederhana terdiri dari tiga lapisan: *input layer*, *hidden layer*, dan *output layer*. Setiap *layer* berisi beberapa neuron, dan *input* dari setiap neuron dari *output layer* sesuai dengan *output* dari *layer* sebelumnya dikalikan dengan matriks bobot diikuti dengan menambahkan matriks bias (Namdari & Li, 2018). *Input layer* menerima *input* (tanpa melakukan operasi apapun) kemudian nilai *input* tersebut diteruskan ke *hidden layer*. Pada *hidden layer*, nilai *input* akan diproses dan dilakukan proses perhitungan hasil fungsi aktivasi pada setiap neuron, lalu hasilnya akan diteruskan ke *layer* selanjutnya. *Output* dari *input layer* akan diterima sebagai *input* bagi *hidden layer*. Begitu pula *output* dari *hidden layer* akan diterima sebagai *input* bagi *output layer*. Kegiatan ini dinamakan *feed forward*. MLP menggunakan *backpropagation* sebagai teknik untuk *training*. *Backpropagation* adalah proses yang dilakukan untuk memperbaharui bobot secara bertahap dari *output layer* ke *hidden layer* berdasarkan *error*. Kemudian *error* tersebut dipropagasi ke lapisan sebelumnya. Kinerja MLP dapat dievaluasi berdasarkan dua ukuran: akurasi dan kesalahan generalisasi. Akurasi merepresentasikan kemampuan belajar model. Kesalahan generalisasi menunjukkan seberapa akurat model mampu memprediksi data yang tidak terlihat (Yu et al., 2020). Banyak ahli menggunakan MLP untuk meramalkan harga saham secara efektif dan terbukti bahwa MLP merupakan model terbaik karena menghasilkan *error* yang kecil. Algoritma model *training* MLP dapat dilihat pada Tabel 3.

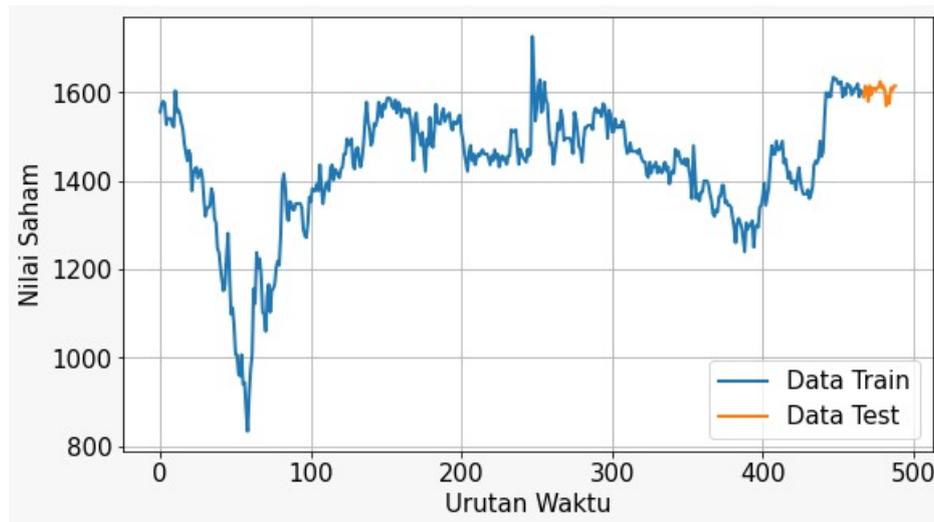
Tabel 3. Algoritma Backpropagation

Input	<ol style="list-style-type: none"> 1. Tetapkan M sebagai banyak iterasi. 2. <i>Training</i> data $X = (\vec{x}_i, y_i)_{i=1}^n$, dengan $\vec{x}_i = (x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, x_i^{(4)}, x_i^{(5)})$
Proses	<ol style="list-style-type: none"> 1. Hitung nilai $o_j = \sigma(\sum_{k=1}^K \vec{x}_k w_{k,j} + \beta_j)$, dengan $\sigma(x) = \max(0, x)$ dan β_j adalah vektor <i>noise</i>. 2. Hitung nilai $v_i = \sigma(\sum_{j=1}^J o_j u_{j,i} + \gamma_i)$. 3. Perbaharui bobot <i>output layer</i> ke <i>hidden layer</i> $\delta_i = (y_i - v_i)v_i(1 - v_i)$ $\Delta u_{j,i} = -\eta(t)\delta_i o_j$ $\Delta \gamma_i = -\eta(t)\delta_i$ 4. Perbaharui bobot <i>hidden layer</i> ke <i>input layer</i> $\varphi_j = \sum_{n=1}^N \delta_i u_{j,i} o_j (1 - o_j)$ $\Delta w_{k,j} = -\eta(t)\varphi_j x_k$ $\Delta \beta_j = -\eta(t)\varphi_j$ 5. Hitung nilai <i>error</i> $L(\theta) = \frac{1}{2} \sum_{n=1}^N (y_i - v_i)^2$ 6. Hitung nilai <i>error</i> yang diturunkan terhadap tiap <i>learning</i> parameter $u_{i,j} = \frac{\delta L(\theta)}{\delta u_{j,i}} = (y_i - v_i)v_i(1 - v_i)o_j$ $w_{j,k} = \frac{\delta L(\theta)}{\delta w_{j,k}} = \sum_{n=1}^N (y_i - v_i)(v_i(1 - v_i)u_{j,i})(o_j(1 - o_j)x_k)$ 7. Ulangi langkah satu sampai dengan enam sebanyak M kali.
Output	arg min $L(\theta)$

3. HASIL DAN PEMBAHASAN

Hasil

Dataset yang digunakan adalah data saham PT. Kalbe Farma Tbk. sebanyak 489 data yang telah divisualisasikan pada Gambar 2. Data tersebut dibagi menjadi dua yaitu 467 data *train* dan 22 data *test*. Selain itu, statistik deskriptif data Saham PT. Kalbe Farma Tbk. dapat dilihat pada Tabel 4 dengan satuan rupiah.



Gambar 2. Visualisasi data Pergerakan Saham PT. Kalbe Farma Tbk.

Tabel 4. Statistik Deskriptif Data Saham PT. Kalbe Farma Tbk.

Statistik	Dataset	Train	Test
Rataan	1439,11	1431,34	1604,09
Median	1459,61	1452,28	1610
Modus	1460,99	1460,99	1615
Minimum	833,38	833,38	1570
Q1	1380	1376,36	1597,50
Q3	1535,69	1525,88	1615
Maksimum	1727,04	1727,38	1625
Standar Deviasi	137,46	135,76	14,53
Skewness	-1,41	-1,45	-1,13
Kurtosis	2,83	2,96	0,47

Hasil

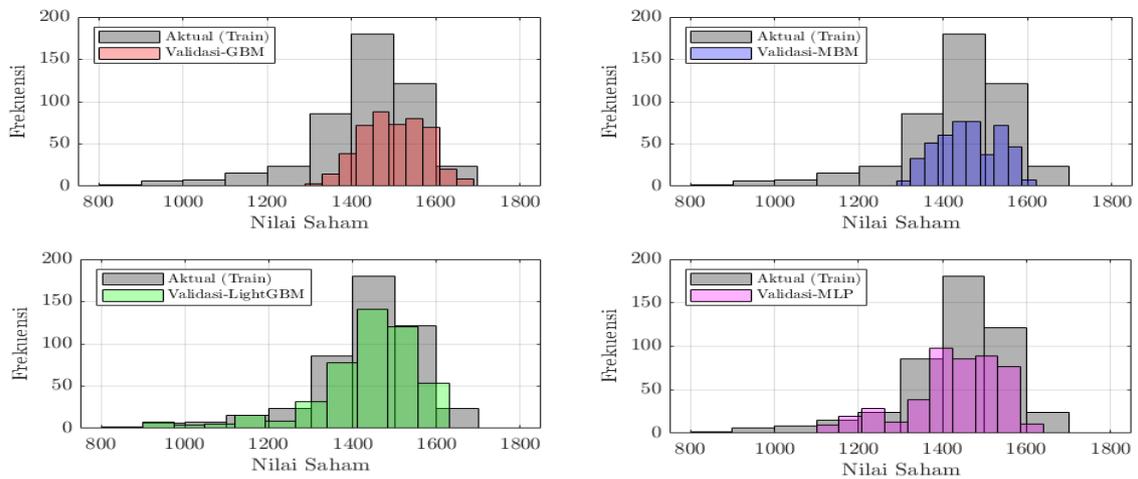
Adapun sebagai tahap persiapan simulasi dari semua model prediksi saham PT. Kalbe Farm Tbk. yaitu melakukan estimasi parameter dengan menggunakan algoritma ARS pada GBM dan MBM, beserta spesifikasi parameter dari LightGBM dan MLP sebagai input pada Tabel 2 dan Tabel 3 disajikan pada Tabel 5.

Tabel 5. Tabel Spesifikasi Parameter untuk GBM, MBM, LightGBM, dan MLP

Model	Parameter	Nilai
GBM	μ	0,00070
	σ	0,01183
	a	0,13059
	b	0,00016
MBM	c	0,00005
	d	0,00005
	σ	0,07118
LightGBM	<i>boosting_type</i>	'gbd'
	<i>num_iteratons</i>	100
	<i>hidden_layer_sizes</i>	(50, 75, 100, 75, 50)
MLP	<i>activation</i>	'relu'
	<i>learning_rate</i>	'adaptive'
	<i>solver</i>	'adam'
	<i>max_iter</i>	5000

Kemudian, simulasi dilakukan dengan menggunakan konfigurasi pada Tabel 5. Simulasi diulang sebanyak 50 kali yang kemudian dirata-rata untuk melihat apakah hasil yang diperoleh sudah konvergen ke suatu nilai. Hasil pertama pada Gambar 3 menunjukkan bahwa GBM dan MBM memberikan wujud

distribusi yang kurang sesuai dengan data aktual. Namun, MBM menunjukkan hasil yang cukup menarik dengan distribusinya tepat berada di persekitaran modus dari data aktual. Dengan kata lain, bentuk distribusi GBM digeser ke kiri. Hasil kedua memberikan informasi bahwa model kecerdasan buatan yang digunakan memiliki distribusi yang hampir sama dengan data aktual.



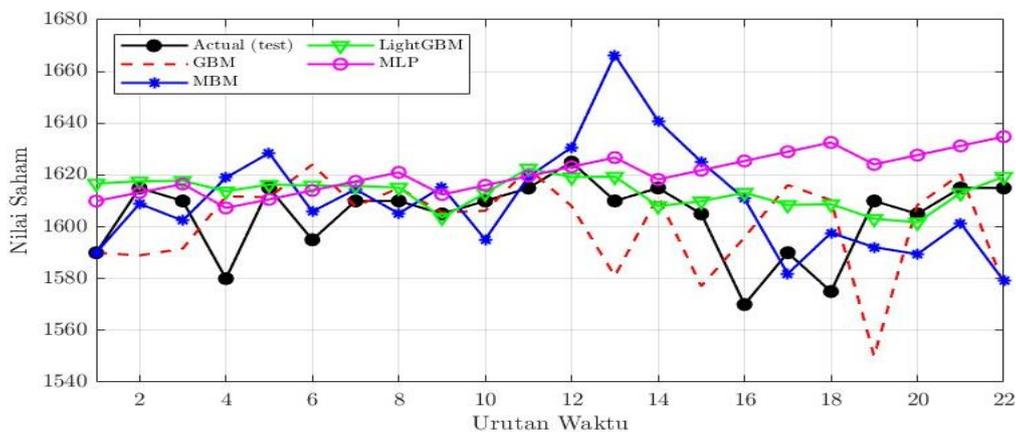
Gambar 3. Hasil Proses Training

Hasil ketiga adalah peninjauan kuantitatif pada performa dari masing-masing model. Metrik performa yang dipilih adalah *Mean Absolute Percentage Error* (MAPE) yang diberikan di Tabel 6. Metrik MAPE digunakan untuk melihat seberapa jauh tingkat kesalahan yang diberikan model terhadap data aktual dalam bentuk persentase yang relatif lebih mudah diinterpretasikan.

Tabel 6. Nilai MAPE pada Proses Training dan Testing

Model	MAPE (%)	
	Proses Training	Proses Testing
GBM	6,89	1,13
MBM	5,80	1,08
LightGBM	1,01*	0,73*
MLP	5,09	1,08

Hasil keempat adalah pemodelan pergerakan harga saham kesehatan dari KLBF.JK lebih cocok apabila dimodelkan menggunakan LightGBM, yang mana memberikan MAPE di bawah angka 1%. Hal ini juga mendukung hasil analisis secara visual pada bagian sebelumnya yang mana hasil prediksi LightGBM mampu memberikan hasil histogram yang relatif mirip dengan data aktualnya. Hal ini dapat dilihat pada Gambar 4.



Gambar 4. Hasil Testing

Hasil kelima adalah informasi pada proses *testing* yang mana semua model memberikan hasil prediksi dengan tingkat kesalahan di kisaran angka 1%. MBM mampu memberikan perbaikan pada GBM yang mana ditunjukkan oleh MAPE yang lebih kecil dibandingkan dengan GBM, yaitu sebesar 1,08%. Meskipun hasil *training* MLP lebih baik daripada MBM, MBM memberikan tingkat MAPE yang lebih kecil pada proses *testing*, dimana hanya berselisih sebesar 0,008%, dibandingkan dengan MLP. Namun, model terbaik pada proses *testing* masih pada LightGBM dengan tingkat kesalahan dibawah 1% yaitu sebesar 0,73%.

Pembahasan

Sebagai aplikasi, terdapat empat model yang digunakan yaitu, Model pertama adalah Model Gerak Brown Geometrik (GBM) dengan estimasi parameternya menggunakan *Maximum Likelihood Estimation* (MLE) (Hull, 2018), Model kedua adalah Gerak Brown Geometrik Termodifikasi (MBM). MBM adalah modifikasi model GBM yang komponen tren dan volatilitasnya diperumum menjadi suatu fungsi yang bergantung pada waktu. Model kedua adalah model yang terinspirasi dari penelitian (Oud & Goard, 2015) yang kemudian dikembangkan pada model opsi (Tong & Liu, 2018). Proses estimasi pada MBM menggunakan algoritma ARS yang terinspirasi dari (Konstantinov et al., 2019) yaitu dengan melakukan *random searching* pada suatu domain tertentu. Penggunaan algoritma ARS pada Tabel 1 digunakan untuk proses estimasi parameter sebab MBM tidak memiliki bentuk yang *closed-form*. Hasil studi dan simulasi, khususnya pada Gambar 3, Gambar 4 dan Tabel 6 memberikan suatu bukti teoritis bahwa model 3/4-volatilitas mampu mengurangi tingkat kesalahan dari model GBM (Oud & Goard, 2015). Model ketiga dan keempat mengadopsi model kecerdasan buatan yaitu, *Light Gradient Boosting Machine* (LightGBM) dan *Multi Layer Perceptron* (MLP). Alasan pemilihan kedua model adalah sifat dominasi (*superiority*) yang dimiliki MLP dan LightGBM dalam melakukan prediksi pergerakan harga (Chen, Zhang, & Lou, 2020; Chhajer et al., 2022; Khang et al., 2021; Rezeki & Ishafit, 2017). Selain itu, hasil *training* dengan LightGBM dan MLP juga tampak memberikan ekor kiri yang mana mengindikasikan bisa menangkap fenomena *skewness*. Namun, model MLP masih belum bisa menghampiri wujud keruncingan atau kurtosis dari data aktual. Adapun pada Gambar 3, tampak posisi modus dari hasil LightGBM menempati daerah interval yang sama dengan data aktual. Dengan demikian, secara visual dapat disimpulkan bahwa hasil simulasi dengan LightGBM lebih mendekati data aktual apabila dibandingkan dengan metode lain. Peneliti lainnya menyatakan bahwa yang memberikan hasil bahwa LightGBM adalah model yang cocok untuk memodelkan harga dan menangkap tren suatu harga aset (Jabeur et al., 2021; Sun et al., 2020).

Apabila keempat model dibandingkan nilai MAPEnya, maka diperoleh LightGBM merupakan model prediksi terbaik untuk menghampiri pergerakan harga saham KLBFIJK. Selain itu, diperoleh informasi menarik bahwa meskipun MAPE model MBM lebih dari model LightGBM, MBM mampu melampaui performa model MLP pada proses *testing*. Dengan demikian, hal ini menunjukkan bahwa MBM mampu disandingkan dengan model kecerdasan buatan dan berpotensi menjadi *benchmark* model prediksi saham.

4. SIMPULAN

Empat jenis model prediksi pergerakan saham PT. Kalbe Farma Tbk. telah dibuat. Model pertama adalah GBM yang memuat komponen tren dan volatilitas konstan. Model kedua adalah MBM yang memodifikasi komponen tren menjadi fungsi waktu dan juga mengubah komponen volatilitas menjadi model 3/4-volatilitas. Hasil studi dan simulasi memberikan informasi bahwa MBM bisa memperbaiki hasil prediksi GBM menjadi lebih akurat yang ditunjukkan dengan nilai MAPE yang lebih rendah. Model ketiga dan keempat adalah model kecerdasan buatan, yaitu LightGBM dan MLP. Kedua model memberikan nilai MAPE *training* yang lebih kecil dibandingkan dengan GBM dan MBM. Akan tetapi, saat proses *testing*, nilai MAPE terkecil diberikan oleh LightGBM sebesar 0,73%. Selain itu, hasil simulasi juga menunjukkan bahwa MAPE *testing* dari MBM dan MLP adalah bernilai sama. Artinya, bisa dilakukan suatu deduksi bahwa MBM memiliki potensi untuk melampaui model kecerdasan buatan.

5. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Riset PPMI KK FMIPA ITB 2022 atas dukungan pendanaan dalam penelitian ini.

6. DAFTAR PUSTAKA

Ashraf, B. N. /. (2021). Stock markets' reaction to COVID-19: Moderating role of national culture. *Finance Research Letters*, 41. <https://doi.org/10.1016/j.frl.2020.101857>.

- Chen, Q., Zhang, W., & Lou, Y. (2020). Forecasting stock prices using a hybrid deep learning model integrating attention mechanism, multi-layer Perceptron, and bidirectional long-short term memory neural network. *IEEE Access*, 8. <https://doi.org/10.1109/access.2020.3004284>.
- Chhajjer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long-short term memory for stock market prediction. *Decision Analytics Journal*, 2. <https://doi.org/10.1016/j.dajour.2021.100015>.
- D., Z., M., H., & Q., J. (2020). Financial Markets Under the Global Pandemic of COVID-19. *Finance Research Letters*, 36. <https://doi.org/10.1016/j.frl.2020.101528>.
- Daul, S., Jaisson, T., & Nagy, A. (2022). Performance attribution of machine learning methods for stock returns prediction. *The Journal of Finance and Data Science*, 8, 86–104. <https://doi.org/10.1016/j.jfds.2022.04.002>.
- Feng, R., Jiang, P., & Volkmer, H. (2021). Geometric brownian motion with affine drift and its time-integral. *Applied Mathematics and Computation*, 395. <https://doi.org/10.1016/j.amc.2020.125874>.
- Huang, L., Leng, H., Li, X., Ren, K., Song, J., & Wang, D. (2021). A data-driven method for hybrid data assimilation with Multilayer Perceptron. *Big Data Research*, 23. <https://doi.org/10.1016/j.bdr.2020.100179>.
- Hull, J. C. (2018). *Options, Futures, and Other Derivatives*. Pearson Education Limited.
- Ichev, R. (2021). Stock price reaction to appointment of a chief health officer during COVID-19. *Journal of Behavioral and Experimental Finance*, 31. <https://doi.org/10.1016/j.jbef.2021.100541>.
- Jabeur, B. S., Khalfaoui, R., & Arfi, B. W. (2021). The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning. *Journal of Environmental Management*, 298. <https://doi.org/10.1016/j.jenvman.2021.113511>.
- Kamara, A. F., Chen, E., & Pan, Z. (2022). An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices. *Information Sciences*, 594, 1–19. <https://doi.org/10.1016/j.ins.2022.02.015>.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... Liu, T. Y. (2017). LightGBM: a highly efficient gradient boosting decision tree [Conference session]. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, USA*.
- Khang, P. Q., Kaczmarczyk, K., Tutak, P., Golec, P., Kuziak, K., Depczyński, R., ... Rot, A. (2021). Machine learning for liquidity prediction on Vietnamese stock market. *Procedia Computer Science*, 192, 3590–3597. <https://doi.org/10.1016/j.procs.2021.09.132>.
- Konstantinov, S., Diveev, A., Balandina, G., & Baryshnikov, A. (2019). Comparative research of random search algorithms and evolutionary algorithms for the optimal control problem of the mobile robot. *Procedia Computer Science*, 150, 462–470. <https://doi.org/10.1016/j.procs.2019.02.080>.
- Na, H., & Kim, S. (2021). Predicting stock prices based on informed traders' activities using deep neural networks. *Economics Letters*, 204. <https://doi.org/10.1016/j.econlet.2021.109917>.
- Namdari, A., & Li, Z. S. (2018). Integrating fundamental and technical analysis of stock market through multi-layer Perceptron. *2018 IEEE Technology and Engineering Management Conference (TEMSCON)*. <https://doi.org/10.1109/temscon.2018.8488440>.
- Oud, M. A., & Goard, J. (2015). Valuation of Options on Oil Futures under the 3/4 Oil Price Model. *International Journal of Theoretical and Applied Finance*, 18(8). <https://doi.org/10.1142/s0219024915500508>.
- Putri, E. R., Tjahjono, V. R., & Imron, C. (2021). A deposit insurance pricing with a multi-state regime-switching volatility. *International Journal of Applied and Computational Mathematics*, 7(6). <https://doi.org/10.1007/s40819-021-01176-2>.
- Qu, Y., Zhang, Z., & Qin, Z. (2020). Wavelet-aided stock forecasting model based on Ensembled machine learning. *2020 The 3rd International Conference on Machine Learning and Machine Intelligence*. <https://doi.org/10.1145/3426826.3426834>.
- Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, 169. <https://doi.org/10.1016/j.eswa.2020.114332>.
- Rezeki, S., & Ishafit, I. (2017). Pengembangan Media Pembelajaran Interaktif untuk Sekolah Menengah Atas Kelas XI pada Pokok Bahasan Momentum. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3(1), 29. <https://doi.org/10.21009/1.03104>.
- Ruslan, S. M., & Mokhtar, K. (2021). Stock market volatility on shipping stock prices: GARCH models approach. *The Journal of Economic Asymmetries*, 24. <https://doi.org/10.1016/j.jeca.2021.e00232>.
- Salisu, A. A., & Vo, X. V. (2020). Predicting stock returns in the presence of COVID-19 pandemic: The role of health news. *International Review of Financial Analysis*, 71. <https://doi.org/10.1016/j.irfa.2020.101546>.

- Sun, X., Liu, M., & Sima, Z. (2020). A novel cryptocurrency price trend forecasting model based on LightGBM. *Finance Research Letters*, 32. <https://doi.org/10.1016/j.frl.2018.12.032>.
- Syaifudin, W. H., & Putri, E. R. (2021). The Application of Model Predictive Control on Stock Portfolio Optimization with prediction based on geometric brownian Motion-Kalman filter. *Journal of Industrial & Management Optimization*. <https://doi.org/10.3934/jimo.2021119>.
- Tao, M., Gao, S., Mao, D., & Huang, H. (2022). Knowledge graph and deep learning combined with a stock price prediction network focusing on related stocks and mutation points. *Journal of King Saud University - Computer and Information Sciences*, 34(7), 4322–4334. <https://doi.org/10.1016/j.jksuci.2022.05.014>.
- Tiwari, S., Bharadwaj, A., & Gupta, S. (2017). Stock price prediction using data analytics. *International Conference on Advances in Computing, Communication and Control (ICAC3)*. <https://doi.org/10.1109/icac3.2017.8318783>.
- Tong, Z., & Liu, A. (2018). Analytical pricing of discrete arithmetic Asian options under generalized CIR process with time change. *International Journal of Financial Engineering*, 5(1). <https://doi.org/10.1142/s2424786318500020>.
- Yu, Z., Qin, L., Chen, Y., & Parmar, M. D. (2020). Stock price forecasting based on LLE-BP neural network model. *Physica A: Statistical Mechanics and Its Applications*, 553. <https://doi.org/10.1016/j.physa.2020.124197>.
- Zhang, X., Ding, Z., Hang, J., & He, Q. (2022). How do stock price indices absorb the COVID-19 pandemic shocks? *The North American Journal of Economics and Finance*, 60. <https://doi.org/10.1016/j.najef.2022.101672>.
- Zhou, F., Zhang, Q., Sornette, D., & Jiang, L. (2019). Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices. *Applied Soft Computing*, 84. <https://doi.org/10.1016/j.asoc.2019.105747>.
- Zolfaghari, M., & Gholami, S. (2021). A hybrid approach of adaptive wavelet transform, long short-term memory and ARIMA-GARCH family models for the stock index prediction. *Expert Systems with Applications*, 182. <https://doi.org/10.1016/j.eswa.2021.115149>.