

Penerapan Least Squares Support Vector Machines (LSSVM) dalam Peramalan Indonesia Composite Index

Andri Triyono¹, Rahmawan Bagus Trianto², Dhika Malita Puspita Arum³

^{1,2,3}Program Studi Ilmu Komputer, Universitas An Nuur, Grobogan, Indonesia, 58111
e-mail: ¹andritriyono1@gmail.com, ²rahmawanbagust@gmail.com, ³dhika.malita11@gmail.com

Submitted Date: April 08th, 2021

Reviewed Date: May 31st, 2021

Revised Date: June 04th, 2021

Accepted Date: June 15th, 2021

Abstract

In the era of very rapidly advancing technology like today, both internet technology and computerization have made various corporate agencies or investors start thinking about the importance of the stock market in their capital division. Previously there were various purchases by the company's capital, such: gold, land, buildings, production machines, but at this time the purchase of capital shares should also start to attract attention and these purchases are legal investments. Various kinds of company shares that are sold can already be seen through the internet and it is very easy and attractive for companies that will make capital purchases, even the model can be chosen for both long-term and short-term capital purchases. This stock price forecasting system using the Least Squares Support Vector Machines (LSSVM) method will be very popular with investors to help determine conclusions for buying shares because it can reduce losses or even make the right decisions so that it will increase profits for investors or companies. Least Squares Support Vector Machines is a simpler model and has been modified from the previous model, namely: Support Vector Machines (SVM) method. Solving linear equations can be solved in a simpler way using LSSVM compared to using SVM. The variable used in the network is the close price variable. The kernel that used for this study is the RBF kernel. This study consists of three phases or stages. The first stage uses 400 historical data rows, second stage uses 800 historical data rows, and the third stage uses 1200 rows of data. This research obtains the best result of accuracy in the third stage. The third stage has the smallest MSE value: 0.00025248 by using 1200 rows of historical data.

Keywords: LSSVM; RBF Kernel; Stock Price; Prediction

Abstrak

Pada masa teknologi maju sangat pesat seperti sekarang ini baik teknologi internet maupun komputerisasi membuat berbagai instansi perusahaan atau investor mulai berpikir pentingnya pasar saham pada divisi permodalannya. Sebelumnya terdapat berbagai pembelian oleh modal perusahaan, seperti: emas, tanah, bangunan, mesin produksi, akan tetapi pada saat sekarang ini pembelian saham modal juga sudah mulai menarik perhatian dan pembelian tersebut adalah pembelian yang sah. Berbagai macam saham perusahaan yang dijual sudah dapat dilihat melalui internet dan itu sangat memudahkan dan menarik bagi perusahaan yang akan melakukan pembelian modal, bahkan modelnya pun bisa dipilih baik investasi pembelian modal dalam jangka Panjang maupun pendek. System peramalan harga saham ini yang menggunakan metode Least Squares Support Vector Machines(LSSVM) akan sangat digemari para investor untuk membantu menentukan kesimpulan untuk melakukan pembelian saham karena dapat mengurangi kerugian atau bahkan dapat mengambil keputusan yang tepat sehingga akan meningkatkan keuntungan bagi investor atau perusahaan. Least Squares Support Vector Machines adalah model yang lebih sederhana dan sudah di modifikasi dari model sebelumnya, yaitu: metode Support Vector Machines (SVM). Pemecahan persamaan linier dapat dipecahkan dengan cara yang lebih sederhana menggunakan LSSVM dibandingkan dengan menggunakan SVM. Variabel yang digunakan dalam jaringan tersebut adalah variable harga close. Kernel yang digunakan untuk penelitian ini adalah kernel RBF. Penelitian ini terdiri dari tiga tahap. Tahap pertama menggunakan 400 baris data histori, 800 data histori untuk tahap kedua, dan 1200 baris data pada tahap ketiga. Penelitian ini mendapatkan hasil akurasi terbaik pada tahap

ketiga. Tahap ketiga memiliki nilai MSE terkecil, yaitu 0.00025248 dengan menggunakan 1200 baris data histori.

Kata Kunci: LSSVM; RBF Kernel; Harga Saham; Peramalan

1. Pendahuluan

Algoritma yang digunakan untuk meramalkan harga saham adalah algoritma deret waktu yang sangat sulit dilakukan jika menggunakan cara konvensional tanpa dibantu dengan teknologi komputerisasi (Kara, Boyacioglu, & Baykan, 2011). Banyak variable-variable, noise, atau proses yang mempengaruhi pergerakan harga saham dikarenakan pergerakan tersebut bersifat non linier dan juga non stasioner sehingga kondisi harga saham akan selalu berubah (Kumar & Thenmozhi, 2012). Bahkan jika terdapat perubahan kebijakan negara, kebijakan perusahaan, keadaan politik, dan keadaan ekonomi, psikologi investor, maka akan membuat kondisi finansial juga ikut berubah (Tan, Quek, & Ng, 2007).

Para stakeholder yang berada di dalam dunia saham akan lebih mudah untuk mengambil keputusan dengan adanya system peramalan dengan keakuratan hasil yang tinggi. Dengan ini resiko para investor akan lebih teratasi dan dapat memaksimalkan keuntungan dari perdagangan saham (Ou & Wang, 2011).

Perkembangan teknologi kecerdasan buatan sekarang ini memunculkan banyak metode-metode peramalan dari algoritma sebelumnya. Sehingga untuk menggunakan algoritma tersebut dapat dilakukan dengan komputerisasi yang jelas lebih mudah menggunakan, lebih cepat, dan lebih akurat. Teknologi komputerisasi jelas dapat mengurangi kesalahan pengguna dari sebuah deret perhitungan yang rumit. Teknologi ini membuat para investor atau pemain modal tertarik untuk menggunakannya dan sangat menjanjikan dengan keakuratan yang lebih tinggi (Tan, Quek, & Ng, 2007). Penelitian tentang peramalan menggunakan kecerdasan buatan sudah dilakukan oleh berbagai peneliti di dunia, antara lain: Penelitian tentang peramalan harga saham di Istanbul pada tahun 2011 yang dilakukan oleh Baykan, Boyacioglu, dan Kara. Pada tahun 2007 juga dilakukan penelitian untuk membandingkan dua metode kecerdasan buatan untuk peramalan, metode tersebut adalah metode SVM dan metode Random Forest, penelitian ini dilakukan oleh Kumar & Thenmozhi. Penelitian perbandingan metode kecerdasan buatan ini juga dilakukan oleh Ou & Wang pada tahun 2009, peneliti membandingkan hingga sepuluh metode kecerdasan buatan untuk dibandingkan tingkat

keakuratannya dalam meramalkan pergerakan harga saham.

2. Landasan Teori Peramalan Saham

Peramalan saham adalah salah satu aktivitas predictive modeling yang mencakup proses pengumpulan data, formulasi model peramalan, dan kemudian pelaksanaan peramalan serta validasi model. Model algoritma untuk meramalkan kejadian masa depan menggunakan berbagai Teknik yang saling berhubungan dan mempengaruhi hasil, diantara Teknik tersebut adalah Teknik mengumpulkan data atau data mining pada data di masa sekarang dan juga data di masa lampau (Mishra, Das, Mausumi, & Mishra, 2011).

Least Squares Support Vector Machines (LSSVM)

LSSVM (Least Squares Support Vector Machines) merupakan sebuah metode yang merupakan hasil modifikasi dari metode Support Vector Machines (SVM) yang dilakukan oleh (Suykens et al, 2002). Pada penelitian Suykens selanjutnya, Suykens berhasil membandingkan LSSVM dengan SVM dalam pemecahan model perhitungan linier dan terbukti LSSVM memiliki tingkat keakuratan yang lebih tinggi dan lebih sederhana dalam menyelesaikan permasalahan perhitungan linier. LSSVM juga memiliki waktu proses yang lebih cepat dibandingkan SVM dalam perhitungan linier pada berskala besar, dengan kata lain LSSVM menggunakan resource computer yang lebih sedikit untuk prosesnya (Mustaffa et al, 2011). Dalam menggunakan kernel RBF, jumlah parameter yang harus dioptimalkan pada LSSVM juga lebih sedikit dibanding SVM, karena pada LSSVM hanya ada 2 variable yang perlu ditentukan, yaitu (γ , σ^2), sedangkan pada SVM memiliki 3 parameter yang harus ditentukan, yaitu (C , σ^2 , ϵ) (Ou & Wang, 2011).

LSSVM juga sama seperti SVM dalam peruntukannya. LSSVM juga digunakan dalam penyelesaian permasalahan regresi dan klasifikasi baik pada kasus linear ataupun non linear. Teknik kernel juga dapat dilakukan di LSSVM pada kasus non linier. Kernel yang dapat digunakan antara lain: MLP, RBF, linier, dan polinomial.

3. Metodologi

Pada figure 1. ditampilkan rancangan eksperimen yang akan dilakukan dalam penelitian ini. Pertama eksperimen dilakukan dengan pengumpulan data yang berupa data harga penutupan dai Indonesia Composite index. Data yang sudah dikumpulkan akan dibagi menjadi tiga data set, yaitu training data set, validation data set, dan testing data set. Feature set yang dihasilkan akan digunakan oleh model LSSVM untuk proses training, validation, dan testing.

Setelah tahap training selesai maka kinerja dari model LSSVM akan dievaluasi menggunakan validation set, kemudian setelah terpilih kinerja yang paling baik maka akan dilakukan proses testing. Pada tahap setelah testing maka hasil tingkat keakuratan bisa didapatkan.

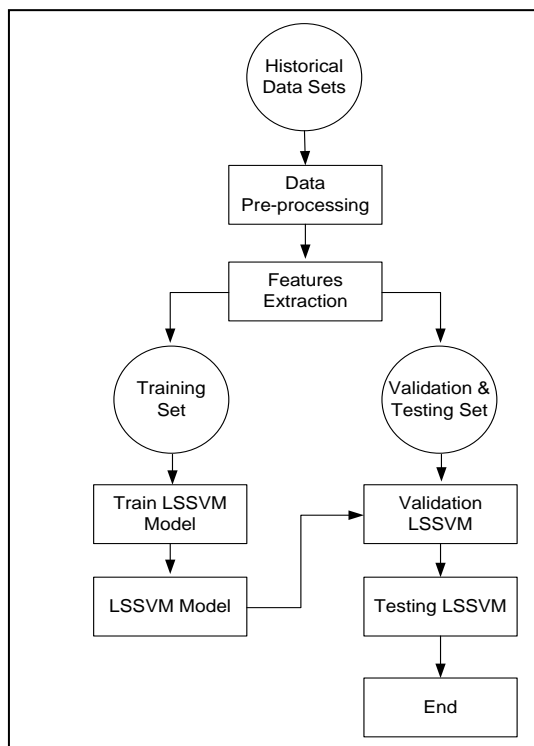


Figure 1. Metodologi yang diusulkan

Pengumpulan Data

Data masa lampau diambil dari tahun 2010 hingga 2015 yang diambil data situs finance.yahoo.com. Data tersebut berupa data harga penutupan per hari dengan total data 1241 baris data yang data tersebut akan kita lakukan proses normalisasi.

Tabel Data

Tabel 1. Sample Data

Tanggal	Close
01/04/2010	25754130

02/04/2010	26052770
03/04/2010	26032970
...	...
31/07/2015	48025290

Data tersebut merupakan data dari situs finance.yahoo.com yang akan digunakan dalam penelitian ini. Total data yang diambil adalah 1241 baris data yang diambil dari tanggal 1 April 2010 hingga 30 Juli 2015. Data yang diambil adalah data harian dimana data tersebut menampilkan pergerakan nilai variable data setiap harinya.

Data yang disiapkan pada penelitian ini akan dibagi menjadi tiga, yaitu data training, data validasi, dan data testing. Besar pembagian data tersebut adalah 70% untuk data training, 15% untuk data validasi, dan 15% untuk data testing. Berikut adalah tabel untuk pembagian data tersebut:

Tabel 2. Tabel pembagian untuk training, validation, dan testing.

Data Set	Periode	Jumlah Observasi
Training	01/04/2010 - 13/03/2013	797
Validation	14/03/2013 - 25/04/2014	228
Testing	28/04/2014 - 31/07/2015	210

4. Hasil dan pembahasan

Penelitian ini dilakukan secara bertahap, tahapan tersebut adalah tahap pembagian data training. Tahap pertama akan dilakukan dengan data training sebanyak 400 baris data, tahap kedua dilakukan dengan 800 baris data, dan tahap ketiga dilakukan dengan 1235 baris data.

Software yang digunakan untuk penelitian ini adalah Matlab 2015a. Konfigurasi untuk pelatihan data dituliskan dengan code:

```

for i = 1:1-20
Xstart(1:20,i) = d(i:i+19)';
    
```

Artinya data yang digunakan untuk iterasi pertama adalah data hari pertama sampai hari ke 20, data itu untuk meramalkan hari 21, dan akan dilakukan perulangan seterusnya sampai data hari terakhir.

Selanjutnya mempersiapkan model untuk peramalan. Pada model ini digunakan kernel RBF, maka code yang dituliskan sebagai berikut:

```

pre=predict({Xtra,Ytra,'f',gam,sig2,'RBF_kernel'},Xstart(:,i),nth);
    
```

Untuk mendapatkan perbandingan hasil peramalan dari data asli dan sekaligus menampilkannya ke dalam figure, maka memanggil fungsi Ymatrix, atau bisa dilihat pada code di bawah ini:

```
Ymatrix = [dataAsli(lag :end) ,
prediksi(lag:end)];
createfigure (Ymatrix);
```

Tahap Pertama

Pada tahap pertama yaitu peramalan menggunakan 400 baris data, maka hasilnya dapat dilihat pada figure di bawah ini:

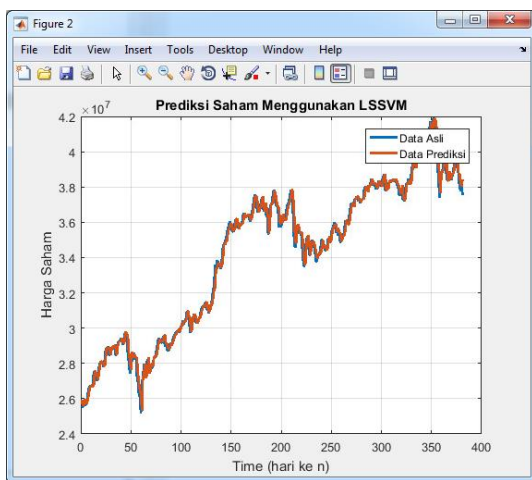


Figure 2. perbandingan pergerakan data hasil peramalan dengan data asli (400 baris data)

Pergerakan hasil peramalan ditunjukkan dengan garis yang berwarna merah, sedangkan data history ditunjukkan dengan garis yang berwarna biru.

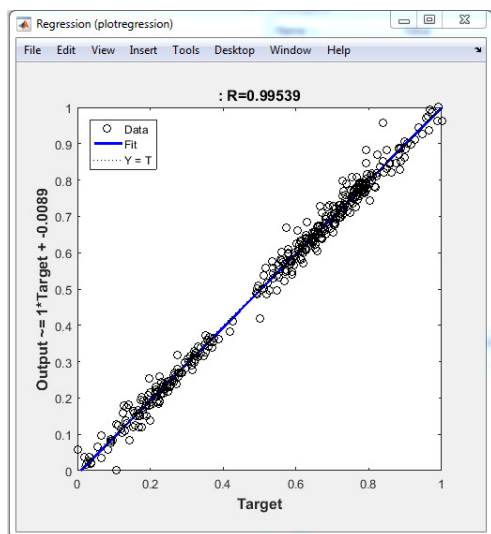


Figure 3. pola regresi terhadap data histori (400 baris data)

Pada figure pola regresi diatas menunjukkan bahwa pengaruh data histori untuk pelatihan cukup tinggi, yaitu sebesar 99.539%.

Untuk menghitung nilai error atau MSE maka data asli dan hasil peramalan harus dinormalisasikan dahulu, maka dituliskan code di bawah ini:

```
dtNor = normalisasi(dataAsli(lag :end), 0, 1);
predNor = normalisasi(prediksi(lag:end), 0, 1);
```

dan untuk menampilkan nilai MSE nya maka dituliskan code:

```
errMse = immse (dtNor ,predNor )
```

maka nilai MSE yang dihasilkan adalah sebesar 6.3964e-04 atau 0.00063964.

Tahap kedua

Pada tahap kedua, yaitu peramalan menggunakan 800 baris data, hasil dapat kita lihat di bawah ini:

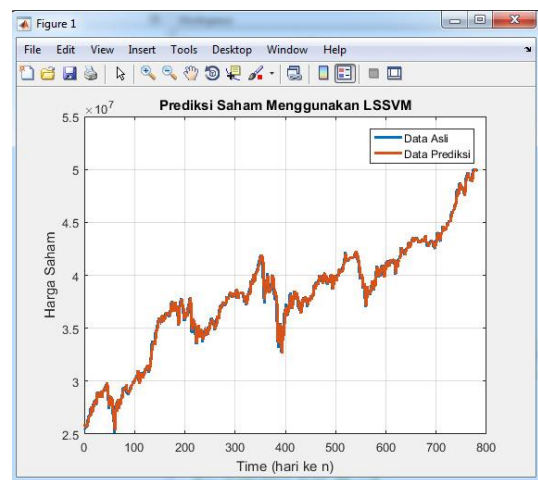


Figure 4. perbandingan pergerakan data hasil peramalan dengan data asli (800 baris data)

Pada figure diatas menunjukkan bahwa garis merah (hasil peramalan) sudah semakin akurat dalam mengikuti pergerakan garis biru (data histori).

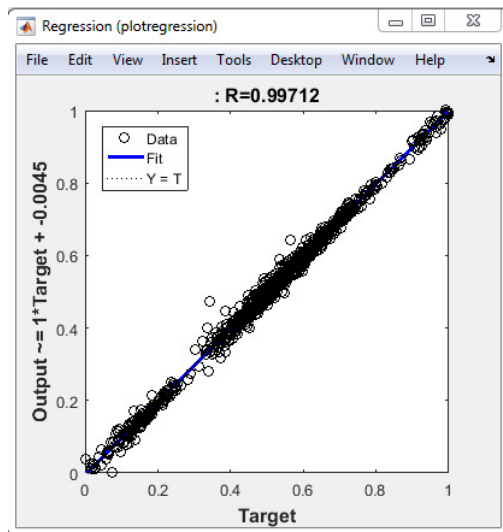


Figure 5. pola regresi terhadap data histori (800 baris data)

Nilai regresi pada peramalan menggunakan 800 baris data adalah sebesar 99.712%. Nilai regresi ini lebih besar disbanding nilai regresi pada peramalan tahap pertama.

Pada peramalan menggunakan 800 baris data ini menghasilkan nilai error (MSE) sebesar $2.8598e-04$ atau sebesar 0.00028598, dan nilai ini juga lebih kecil disbanding peramalan pada tahap pertama.

Tahap ketiga

Pada tahap ketiga, yaitu peramalan menggunakan 1200 baris data, hasil dapat kita lihat di bawah ini:

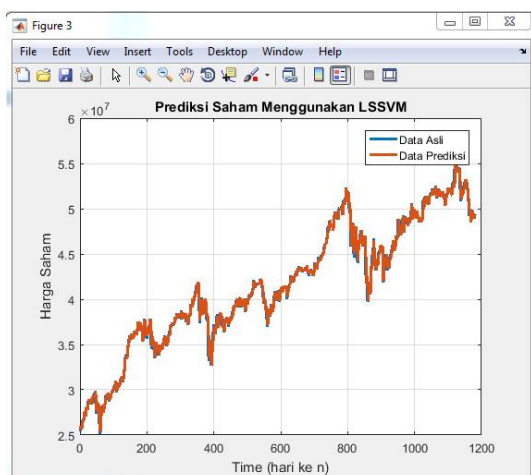


Figure 6. perbandingan pergerakan data hasil peramalan dengan data asli (1200 baris data)

Pada figure diatas garis merah(hasil peramalan) semakin mengikuti garis biru(data histori). Terlihat garis merah semakin banyak dapat menutupi garis biru, maka bisa terlihat penelitian tahap 3 ini

menghasilkan peramalan yang lebih akurat.

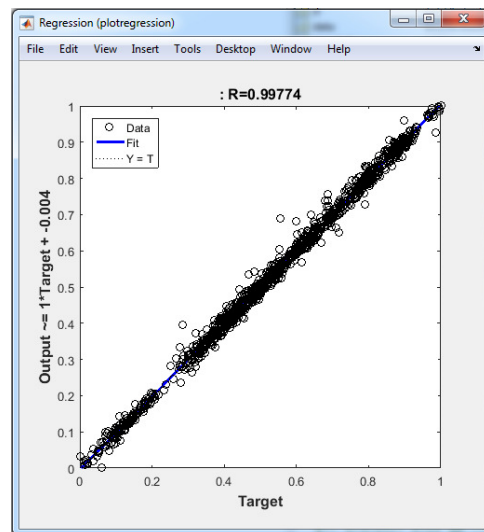


Figure 7. pola regresi terhadap data histori (1200 baris data)

Pada pola regresi diatas terlihat semakin mendekati data histori. Nilai regresi pada penelitian tahap ketiga ini adalah 99.774%.

Pada peramalan menggunakan 1200 baris data ini menghasilkan nilai error (MSE) sebesar $2.5248e-04$ atau sebesar 0.00025248, dan nilai ini juga lebih kecil disbanding peramalan pada tahap pertama dan kedua.

Hasil penelitian dari ketiga tahap tersebut bisa dilihat pada tabel di bawah ini:

Table 2. Hasil Penelitian

Tahap	Regresi	MSE
1	99.539%.	0.00063964
2	99.712%.	0.00028598
3	99.774%.	0.00025248

5. Kesimpulan

Dengan melihat tabel hasil penelitian pada pembahasan diatas, perubahan nilai MSE terlihat jauh antara penelitian yang menggunakan data histori sebanyak 400 baris dan 800 baris, yaitu 0.00063964 dan 0.00028598, dan perubahan nilai MSE antara 800 baris dan 1200 baris perubahan nilai MSE tidak terlalu jauh, yaitu 0.00028598 dan 0.00025248.

Hasil dari penelitian ini, diperoleh bahwa semakin banyak data histori, maka hasil peramalan juga semakin akurat. Hal ini dibuktikan dengan melihat nilai error (MSE) yang semakin kecil ketika data histori diperbanyak. Metode LSSVM memiliki tingkat akurasi terbaik atau nilai MSE terkecil pada pelatihan data histori sebanyak 1200, yaitu 0.00025248.

Referensi

- Astudillo, G., Carrasco, R., Fernández-Campusano, C., & Chacón, M. (2020). Copper Price Prediction Using Support Vector Regression Technique. *Applied Sciences; Basel*, 6648.
- Chenglin, X., Weili, X., & Jijiao, J. (2020). Stock price forecast based on combined model of ARIMA-LS-SVM. *Neural Computing & Applications; Heidelberg*, 5379-5388.
- Desfiandi, A., Desfiandi, A., & Ali, H. (2017). Composite Stock Price Index (IHSG) Macro Factor in Investment in Stock (Equity Funds). *International Journal of Economics and Financial Issues; Mersin*, 534-536.
- Farhadi, M., & Mollayi, N. (2019). Application of the Least Square Support Vector Machine for point-to-point forecasting of the PV Power. *International Journal of Electrical and Computer Engineering; Yogyakarta*, 2205-2211.
- Guo, T., He, W., Jiang, Z., Chu, X., Malekian, R., & Li, Z. (2019). An Improved LSSVM Model for Intelligent Prediction of the Daily Water Level. *Energies; Basel*, 112.
- Guo, W., & Zhao, Z. (2017). A Novel Hybrid BND-FOA-LSSVM Model for Electricity Price Forecasting. *MDPI AG*, 120.
- Huang, C.-H., Yang, F.-H., & Lee, C.-P. (2018). The strategy of investment in the stock market using modified support vector regression model. *Scientia Iranica. Transaction E, Industrial Engineering; Tehran*, 1629-1640.
- Kara, Y., Boyacioglu, M. A., & Baykan, O. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications*, 5311-5319.
- Kumar, M., & Thenmozhi, M. (2012). Stock Index Return Forecasting and Trading Strategy Using Hybrid ARIMA-Neural Network Model. *International Journal of Financial Management; New Delhi*, 1-14.
- Luxianto, R., Arief, U., & Prasetyo, M. B. (2020). Day-of-the-Week Effect and Investors' Psychological Mood Testing in a Highly Mispriced Capital Market. *Journal of Indonesian Economy and Business : JIEB; Yogyakarta*, 257-269.
- Mishra, D., Das, A. K., Mausumi, M., & Mishra, S. (2011). Predictive Data Mining: Promising Future and Applications. *International Journal of Computer and Communication Technology*, 3.
- Murti, W. (2017). The Effects of Gold Prices, the Hang Seng Index, and the Dow Jones Index on the Composite Stock Price Index (CSPI) in Indonesia. *International Journal of Economic Perspectives; Mersin*, 258-266.
- Ou, P., & Wang, H. (2011). Modeling and forecasting stock market volatility by Gaussian processes based on GARCH, EGARCH and GJR models. *Proceedings of the World Congress on Engineering* (pp. 338-342). London, U.K: WCE.
- Pai, P.-F., Ling-Chuang, H., & Kuo-Ping, L. (2018). Using Internet Search Trends and Historical Trading Data for Predicting Stock Markets by the Least Squares Support Vector Regression Model. *Computational Intelligence and Neuroscience : CIN; New York*, 1155.
- Tan, T. Z., Quek, C., & Ng, G. S. (2007). Biological Brain-Inspired Genetic Complementary Learning For Stock Market And Bank Failure Prediction. *Computational Intelligence*, 236-261.