

KAJIAN KOMPARASI PENERAPAN ALGORITMA *SUPPORT VECTOR MACHINE (SVM)* DAN *MULTILAYER PERCEPTRON (MLP)* DALAM PREDIKSI INDEKS SAHAM SEKTOR PERBANKAN: STUDI KASUS SAHAM LQ45 IDX BANK BCA

Taufik Hidayatulloh

Program Studi Manajemen Informatika

AMIK BSI Jakarta

Jl. RS Fatmawati No 24 Pondok Labu 12450, Indonesia

e-mail: taufik.tho@bsi.ac.id

Abstrak- *Prediksi harga saham merupakan isu penting dalam investasi, selalu menarik perhatian investor publik dan akademisi untuk penelitian. Akan tetapi, fakta harga saham yang berfluktuasi dan tantangan bagi para analis untuk diprediksi. Nilai dari pergerakan saham bersifat dinamis dan tidak linier sehingga diperlukan metode khusus untuk model secara matematis. Dalam studi komparasi ini dilakukan untuk membandingkan algoritma Multilayer Perceptron (MLP) dan Support Vector Machine (SVM) untuk meramalkan/memprediksi indeks saham harian LQ45 khusus sektor perbankan BEI diambil dari delapan posisi teratas. Hasil pembobotan Bank Central Asia Tbk menempati posisi teratas dan selalu ada. Variabel prediksi dibedakan dalam tiga model, yaitu input 1 lag, 6 lag dan 10 lag. Dari berbagai model yang dapat dibentuk dengan arsitektur jaringan, fungsi aktivasi, jaringan metode pembelajaran yang berbeda pada Multilayer Perceptron (MLP) dan jenis SVR, fungsi kernel Support Vector Regression (SVR) persiapan dan pelaksanaan data dengan menghapus kecenderungan data varian time series (rentet waktu) diikuti oleh stabilisasi atau tidak, diperoleh 48 model untuk MLP dan 16 untuk SVR pada masing-masing model prediksi. Hasil empiris menunjukkan bahwa model Support Vector Regression (SVR) dengan 10 prediksi variabel lag memberikan akurasi peramalan yang terbaik dibandingkan dengan SVR model dengan variabel tertinggal dan prediksi 6 MLP model kedua model 6 lag dan 10 lag.*

Kata Kunci: *Peramalan, prediksi, harga saham, Neural Network, MLP, SVM*

I. PENDAHULUAN

Saham (*stock*) merupakan salah satu instrumen pasar keuangan yang paling populer. Saham dapat didefinisikan sebagai tanda penyertaan modal seseorang atau pihak (badan usaha) dalam suatu perusahaan atau perseroan terbatas [1].

Prediksi harga saham merupakan salah satu masalah penting di bidang keuangan. Dan bursa prediksi harga yang bervariasi tergantung pada waktu dan informasi masa lalu. Hal ini digunakan untuk menentukan nilai masa depan dari suatu saham perusahaan atau instrumen keuangan lainnya yang diperdagangkan di bursa keuangan [2]. Prediksi pasar saham adalah hal penting dan sangat menarik karena keberhasilan prediksi harga saham dapat menjanjikan keuntungan yang menarik. Isu tentang prediksi harga saham sudah banyak diangkat dalam makalah penelitian [3]. Akurasi hasil prediksi indeks pasar saham digunakan untuk banyak alasan, utamanya adalah untuk kebutuhan bagi investor untuk melindungi nilai terhadap risiko pasar yang potensial, dan kesempatan untuk spekulator pasar dan arbitrase untuk membuat keuntungan dari indeks perdagangan.

Metode prediksi rentet waktu seperti *Support Vector Machine (SVM)*, *Neural Network (NN)* [4] dan

Grey Model (GM) (Wu, 2007) diusulkan oleh banyak peneliti [5].

MLP (Multilayer Perceptron) merupakan salah satu metode dari *Neural Network* sangat cocok untuk menyelesaikan masalah yang tidak linear dan non deterministik [6]. MLP telah diterapkan dengan sukses untuk menyelesaikan masalah-masalah yang sulit dan beragam dengan melatihnya menggunakan algoritma propagasi balik dari kesalahan atau *error backpropagation (EBP)* [7]. Sedangkan kelebihan *Support Vector Machine (SVM)* adalah SVM tidak mengalami *overfitting* karena training perlu dilakukan sekali saja dan mendapatkan solusi optimal [8].

Penelitian ini diawali dengan mencari sifat multifraktal pada return saham objek penelitian dengan analisis *rescaled range* (untuk mendapatkan eksponen hurst) untuk mengetahui apakah data return tersebut bersifat acak atau terdapat pengulangan trend sehingga dapat dilakukan analisis teknikal. Selanjutnya akan dilakukan prediksi terhadap return saham tersebut dengan metode SVM (*Support Vector Machines*) dan MLP (*Multilayer Perceptron*) untuk kemudian akan dilakukan komparasi metode mana yang memiliki kesalahan lebih kecil dalam memprediksi indeks LQ45.

Indeks LQ45 adalah nilai kapitalisasi pasar dari 45 saham yang paling likuid dan memiliki nilai kapitalisasi yang besar hal itu merupakan indikator likuidasi. Tujuan indeks LQ45 adalah sebagai pelengkap Indeks Harga Saham Gabungan (IHSG) dan khususnya untuk menyediakan sarana yang obyektif dan terpercaya bagi analisis keuangan, manajer investasi, investor dan pemerhati pasar modal lainnya dalam memonitor pergerakan harga dari saham-saham yang aktif diperdagangkan. Penilaian kinerja secara *raw performance* dan *risk performance*, ditemukan bahwa kinerja Reksadana Indeks LQ45 lebih baik dari pada kinerja Reksadana saham (secara keseluruhan) di Pasar Modal Indonesia [9].

Beberapa penelitian yang telah dilakukan dipasar modal Indonesia, sebagian besar hanya melakukan kajian yang berkaitan dengan analisis fundamental saja [2] karena sebagian besar data diperoleh dari laporan keuangan perusahaan yang dipublikasikan, tetapi sangat sedikit sekali yang melakukan kajian terhadap analisis teknikal. Berbagai macam metode dapat digunakan untuk melakukan analisis teknikal pada pasar modal. Oleh karena itu, perlu diketahui performa prediksi tiap metode agar prediksi dapat dilakukan dengan lebih baik.

II. LANDASAN TEORI

Analisa *time series* adalah salah satu metode peramalan secara matematis dengan menggunakan waktu sebagai acuan, kemudian membuat prediksi dengan menggunakan ekstrapolasi berdasarkan waktu untuk pola-pola tersebut. [10]

Data mining adalah sebuah proses, yang mana dalam melakukan prosesnya harus sesuai dengan prosedur dari proses, yaitu CRISP-DM (*Cross-Industry Standard Process for Data Mining*), yang terdiri dari keseluruhan proses, preprosesing data, pembentukan model, model evaluasi, dan tahap akhir penyebaran model [11].

Metode penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif untuk menganalisa data rentet waktu (*time series*). Dengan studi kasus harga penutupan saham harian LQ45 IDX khusus sektor perbankan, dalam hal ini adalah Bank Central Asia Tbk (BBCA.JK).

2.1. Jenis Penelitian

Jenis penelitian dalam penelitian ini sebagai berikut:

a. Penelitian Eksperimental

Penelitian eksperimental merupakan penelitian yang bersifat uji coba, memanipulasi dan mempengaruhi hal-hal yang terkait dengan seluruh variabel atau atribut.

b. Penelitian Perbandingan atau Studi Komparasi

Penelitian yang dilakukan dengan cara membandingkan antara dua macam algoritma

yaitu *Multilayer Perceptron (MLP)* dan *Support Vector Machine (SVM)*.

Kegiatan penelitian ini melalui beberapa tahap dalam pengembangannya yaitu:

1. Studi Pendahuluan

Kegiatan yang dilakukan pada saat studi pendahuluan yaitu mengumpulkan materi-materi kepustakaan yang berhubungan dengan pengambilan judul. Kemudian langkah selanjutnya yaitu survey mengenai prediksi indeks saham. Dari penelitian tersebut akan di dapat beberapa kriteria yang digunakan untuk penelitian lebih lanjut.

2. Data

Setelah dilakukannya survey maka didapatkan data yang akan digunakan dalam penelitian ini, yang terdiri dari beberapa atribut yaitu data-data masukan berupa indeks harga saham harian selama periode Juli 2009 sampai dengan Agustus 2013 khusus sektor perbankan.

3. Pengolahan Data

Setelah mendapatkan data. Maka data diolah menggunakan metode data mining, yaitu pertama kali menggunakan Microsoft Excel.

4. Laporan

Setelah tahap studi pendahuluan, data, dan pengolahan data tersebut dilakukan maka disusunlah laporan penelitian ini kedalam karya ilmiah dalam bentuk jurnal.

2.2. Metode Pengumpulan Data

Data sampel dalam penelitian ini berasal dari data sekunder, dan merupakan data runtun waktu, berupa harga penutupan saham harian Bank Central Asia Tbk (BBCA.JK). Pengumpulan data saham dilakukan dengan cara mendownload data harga penutupan saham harian dari situs www.finance.yahoo.com dan data yang diperoleh berjumlah 1024 record.

2.3. Ruang Lingkup

Sejumlah permasalahan yang dibahas dalam usulan penelitian ini akan dibatasi ruang lingkup pembahasannya, antara lain:

- Data-data masukan berupa indeks harga saham perbulan selama periode Juli 2009 sampai dengan Agustus 2013 khusus sektor perbankan.
- Apakah algoritma SVM (*Support Vector Machines*) dan MLP (*Multilayer Perceptron*) dapat digunakan sebagai metode pengambilan keputusan pemecahan masalah dalam keakuratan prediksi indeks saham?

III. PEMBAHASAN

Obyek yang akan diteliti adalah data harga saham penutupan harian Bank Central Asia Tbk (BBCA.JK), selama kurun waktu 01 Juli 2009 sampai dengan 01 Agustus 2013. Data sampel tidak mengambil data harga saham periode tahun

sebelumnya yaitu tahun 2008 sampai pertengahan tahun 2009, dikarenakan hal tersebut untuk menghindari data tidak wajar, karena pada tahun 2008-2009 terjadi krisis ekonomi global, yang mengakibatkan harga saham berjatuh.

Data sampel akan dibagi menjadi data training dan data validasi. Sebagian besar data akan digunakan sebagai data training, dan sisanya untuk validasi terhadap *rule*. Data validasi akan diambil dari beberapa data terakhir transaksi.

a. Instrumen Penelitian

Dalam penelitian ini, pelaksanaan pengolahan data mining, akan menggunakan perangkat lunak berupa *Microsoft Office Excel* dan *file csv (Comma Separated Value)* sebagai tempat penyimpanan data harga saham, dan DTREG sebagai tools pemroses *data mining*. DTREG adalah *software* untuk pemodelan prediktif dan peramalan. DTREG dikembangkan oleh Philip H. Sherrod. Banyak model prediksi yang ditawarkan oleh DTREG. DTREG membangun pohon keputusan klasifikasi dan regresi, *Neural Networks, Support Vector Machine (SVM), GMDH Polynomial Networks, Gene Expression Programs, K-Means Clustering, Discriminant Analysis* dan *Logistic Regression* yang menggambarkan hubungan data dan dapat digunakan untuk memprediksi nilai-nilai untuk pengamatan masa depan. DTREG juga memiliki dukungan penuh untuk analisis *time series*.

b. Metode Analisis dan Pengujian

Dalam Penelitian ini, data yang telah diperoleh yaitu berupa harga saham akan dibagi menjadi data training dan data validasi. Data training dan data validasi diproses dengan metode Jaringan Syaraf Tiruan, model *Multi layer Perceptron (MLP)* dan *Support Vector Machine (SVM)* yang dimainkan dengan *Support Vector Regression (SVR)*. Variabel prediksi akan dibedakan dalam dua model, yaitu 6 *lag* dan 10 *lag*. Balkin [10] merekomendasikan jumlah *lag* pada *Neural Network* untuk prediksi data harian adalah *lag* 1 sampai dengan *lag* 6. Namun dalam eksperimen ini akan dicoba untuk menambahkan juga *lag* 10 sebagai variabel prediksi untuk melihat bagaimanakah hasilnya, apakah hasil *lag* 6 lebih baik dari *lag* 10. Kemudian hasil pemrosesan data mining akan dianalisa, dengan membandingkan tingkat *Error* dari pemrosesan data validasi terhadap algoritma *Multi layer Perceptron (MLP)* dan *Support Vector Regression (SVR)* pada masing-masing model prediksi. Tingkat *Error* yang digunakan berupa: *Maximum Error (ME)*, *Root Mean Squared Error (RMSE)*, *Mean Squared Error (MSE)*, *Mean Absolute Error (MAE)*, dan *Mean Absolute Percentage Error (MAPE)*. Keputusan ditekankan pada nilai *Root Mean Squared Error (RMSE)*. Setelah itu hasil dari eksperimen model prediksi dengan *lag* 6 dan *lag* 10 akan dibandingkan. Algoritma dengan tingkat *Error* terkecil adalah algoritma yang dianggap paling akurat.

Prosiding SNIT 2014: Hal. 264

Setelah dirancang dan dibuat maka tahap selanjutnya yaitu hasil pengujian.

3.1. Perancangan Eksperimen

Dalam penelitian ini dilakukan dalam dua *setting* eksperimen, yaitu dengan menggunakan dua variabel prediksi yang berbeda. Variabel prediksi sebanyak 6 *lag* dan 10 *lag* akan diterapkan dalam algoritma *Multi layer Perceptron (MLP)* dan *Support Vector Machine (SVM)* yang dimainkan dengan *Support Vector Regression (SVR)*.

Penelitian ini menggunakan dua model prediksi, yaitu *lag* 6 dan *lag* 10, untuk kemudian dibandingkan hasilnya, manakah yang bisa memberikan hasil prediksi dengan *Error* yang lebih kecil. Variabel *input* yang digunakan adalah *lag* hari dari data observasi, dalam hal ini harga saham untuk memprediksi harga saham berikutnya. Apabila dirumuskan, maka rumus prediksinya sebagai berikut:

$$Y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-6}) \quad (4.2)$$

dan

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-10}) \quad (4.3)$$

dimana Y_t adalah data observasi pada hari ke - t

3.2. Data Eksperimen

Dalam penelitian ini data yang digunakan adalah data runtun waktu selama kurun waktu 01 Juli 2009 sampai dengan 01 Agustus 2013. Data sampel tidak mengambil data harga saham periode tahun sebelumnya yaitu tahun 2008 sampai pertengahan tahun 2009, dikarenakan hal tersebut untuk menghindari data tidak wajar, karena pada tahun 2008-2009 terjadi krisis ekonomi global, yang mengakibatkan harga saham berjatuh yang memungkinkan mengacaukan pola yang akan terbentuk. Data yang diperoleh berjumlah 1024 record.

Tabel 1. Statistik Data Sampel

| | |
|--------------------|----------|
| Jumlah Data Sampel | 1024 |
| Mean | 9722.197 |
| Median | 9300 |
| Maximum | 16100 |
| Minimum | 5050 |
| Standard Deviasi | 2370.108 |
| Skewness | 0.629266 |
| Kurtosis | 0.027535 |

3.3. Data Pre-Processing

Agar data siap untuk diolah dalam data mining [11], maka sebelum digunakan dalam eksperimen, data sampel akan dilakukan beberapa proses. Data saham diambil dari Pengumuman PT Bursa Efek Jakarta mengenai saham emiten yang masuk dalam

perhitungan Indeks LQ45 mulai periode Februari 2006 hingga Agustus 2013. Setelah dilakukan penyaringan khusus sektor perbankan, kemudian dilakukan pembobotan untuk mengetahui perusahaan perbankan yang aktif dalam periode 10 tahun terakhir.

Data harga saham yang didownload dari situs finance.yahoo.com, tersimpan dalam file csv (*Comma Separated Value*) dengan kolom: Tanggal, harga pembukaan, harga tertinggi, harga terendah, harga

penutupan, *volume*, *adjustment* penutupan. Sebelum diproses, data sampel diurutkan berdasarkan waktu secara *ascending*. Kolom-kolom yang tidak digunakan didrop, sehingga tinggal satu kolom yaitu harga penutupan saham harian. Setelah data sampel yang telah dibersihkan, maka akan dibagi menjadi data training dan validasi.

Tabel 2. Data Indeks Saham LQ45 Periode Februari 2006 – Agustus 2013

| NO. | 2006 | | 2007 | | 2008 | | 2009 | | Rank |
|-----|------|-------|------|-------|------|-------|------|-------|------|
| | Feb | Agust | Feb | Agust | Feb | Agust | Feb | Agust | |
| 1 | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | 8 |
| 2 | BBRI | BBRI | BBRI | BBKP | BBNI | BBNI | BNGA | BBNI | 7 |
| 3 | BDMN | BDMN | BDMN | BBRI | BBRI | BBRI | BDMN | BBRI | 6 |
| 4 | BMRI | BMRI | BMRI | BDMN | BDMN | BDMN | BNII | BDMN | 5 |
| 5 | BNGA | BNGA | BNGA | BMRI | BMRI | BMRI | BMRI | BMRI | 4 |
| 6 | BNII | BNII | BNII | BNGA | BNGA | BNGA | BBNI | | 3 |
| 7 | BNLI | - | | BNII | BNII | BNII | | - | 2 |
| 8 | - | - | - | | - | - | BBRI | - | 1 |
| NO. | 2010 | | 2011 | | 2012 | | 2013 | | Rank |
| | Feb | Agust | Feb | Agust | Feb | Agust | Feb | Agust | |
| 1 | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | BBCA | 8 |
| 2 | BBNI | BBNI | BBKP | BBNI | BBNI | BBNI | BBNI | BBNI | 7 |
| 3 | BBRI | BBRI | BBNI | BBRI | BBRI | BBRI | BBRI | BBRI | 6 |
| 4 | BDMN | BBTN | BBRI | BBTN | BBTN | BDMN | BBTN | BBTN | 5 |
| 5 | BMRI | BDMN | BBTN | BDMN | BDMN | | BDMN | BDMN | 4 |
| 6 | - | BMRI | BDMN | | | BMRI | BMRI | BMRI | 3 |
| 7 | - | - | | BMRI | BMRI | - | - | - | 2 |
| 8 | - | - | BMRI | - | - | - | - | - | 1 |

Setelah dilakukan penyaringan khusus sektor perbankan, kemudian dilakukan pembobotan untuk mengetahui perusahaan perbankan yang aktif dalam periode 8 tahun terakhir

Tabel 3. Hasil Pembobotan Saham Perbankan LQ45 Periode Februari 2006-Agustus 2013

| No | Emiten | Nama Bank | Jml. Muncul | Total | Bobot |
|----|--------|-----------------------|-------------|-------|-------|
| 1 | BBCA | Bank Central Asia Tbk | 16 | 128 | 8,00 |

| | | | | | |
|----|------|--|----|----|------|
| 2 | BBNI | Bank Negara Indonesia (Persero) Tbk | 12 | 72 | 6,00 |
| 3 | BBRI | Bank Rakyat Indonesia (Persero) Tbk | 16 | 87 | 5,44 |
| 4 | PNBN | Bank Pan Indonesia Tbk | 4 | 12 | 3,00 |
| 5 | BDMN | Bank Danamon Indonesia Tbk | 16 | 73 | 4,56 |
| 6 | BBTN | Bank Tabungan Negara (Persero) Tbk | 6 | 24 | 4,00 |
| 7 | BBKP | Bank Bukopin Tbk | 2 | 14 | 7,00 |
| 8 | BNGA | Bank Niaga Tbk | 6 | 21 | 3,50 |
| 9 | BMRI | Bank Mandiri (Persero) Tbk | 16 | 51 | 3,19 |
| 10 | BNII | Bank International Indonesia Tbk | 7 | 20 | 2,86 |
| 11 | BNLI | Bank Permata Tbk | 1 | 2 | 2,00 |
| 12 | BJBR | Bank Pembangunan Daerah Jabar & Banten Tbk | 4 | 8 | 2,00 |

Dari hasil pembobotan diambil perbankan yang paling aktif selama 8 tahun terakhir dan yang menempati posisi peringkat teratas, yaitu Bank Central Asia (BBCA.JK).

Obyek yang akan diteliti adalah data harga saham penutupan harian Bank Central Asia Tbk (BBCA.JK), selama kurun waktu 01 Juli 2009 sampai dengan 01 Agustus 2013. Data sampel tidak mengambil data harga saham periode tahun sebelumnya yaitu tahun 2008 sampai pertengahan tahun 2009, dikarenakan hal tersebut untuk menghindari data tidak wajar, karena pada tahun 2008-2009 terjadi krisis ekonomi global, yang mengakibatkan harga saham berjatuh.

3.4. Perancangan Arsitektur *Multilayer Perceptron* (MLP)

A. Jumlah Neuron Input

Dalam desain topologi jaringan *Neural Network*, beberapa hal harus ditentukan terlebih dahulu, salah satunya adalah menentukan berapa jumlah neuron input yang akan digunakan dalam jaringan. Jumlah neuron input mengikuti jumlah variabel input yang digunakan. Pada eksperimen ini variabel prediksi berupa *lag* 6 dan *lag* 10 akan digunakan dalam jaringan *Neural Network*.

B. Jumlah Hidden Layer

Penggunaan satu *hidden layer* pada banyak kasus cukup memadai untuk melatih jaringan, namun pada kasus yang lebih kompleks bisa membutuhkan lebih dari satu *hidden layer*. Penggunaan dua *hidden layer* jarang meningkatkan performa model, dan tidak ada alasan teoritis untuk menggunakan lebih dari dua lapisan tersembunyi. DTREG memungkinkan untuk penggunaan 1 sampai 2 *hidden layer*. Pada penelitian ini akan digunakan 1 dan 2 *hidden layer* dalam

topologi jaringan, untuk mendapatkan hasil prediksi terbaik.

C. Jumlah Neuron Hidden Layer

Pada eksperimen ini, penentuan jumlah neuron pada *hidden layer* akan menggunakan fitur otomatis dari DTREG, yaitu pada *hidden layer* pertama. Untuk jumlah neuron pada *hidden layer* kedua akan dilakukan dengan *random*, yaitu mencoba-coba dengan menambahkan beberapa neuron untuk kemudian diamati hasilnya, jumlah neuron pada hasil eksperimen dengan *Error* terkecil yang akan digunakan.

D. Jumlah Output

Pada penelitian ini *output* yang dihasilkan jaringan berjumlah satu dengan jumlah neuron pun satu, karena eksperimen ini merupakan analisa regresi dengan variabel target kontinu. *Output* berupa harga saham hasil prediksi dari variabel *input* data observasi 6 *lag* dan 10 *lag* hari sebelumnya.

E. Fungsi Aktivasi

Pada penelitian ini akan menggunakan fungsi aktivasi yang tersedia pada DTREG, yaitu fungsi logistik (*sigmoid*) dan *linier*. Baik fungsi logistik (*sigmoid*) maupun linier akan diterapkan pada lapisan *hidden* dan lapisan *output*.

F. Metode Pelatihan

Pada penelitian ini akan menggunakan baik *traditional* maupun *scaled gradient conjugate* dalam melatih jaringan untuk menemukan bobot optimal.

G. Model *Multilayer Perceptron* (MLP)

Dari kombinasi *pra preparing* data *time series* oleh DTREG, penerapan jumlah *hidden layer*, fungsi aktivasi serta metode training yang digunakan maka dapat dibuat sebanyak 48 model yang nantinya akan diproses, dan dilihat bagaimana hasil peramalannya.

3.5. Perancangan Struktur *Support Vector Regression* (SVR)

A. Variabel Prediksi

Penelitian ini akan dilakukan dengan variabel prediksi sebanyak 6 *lag*. Selain itu untuk melihat bagaimana pengaruh jumlah penggunaan variabel prediksi, pada eksperimen ini akan meneliti juga penggunaan variabel prediksi sebanyak 10 *lag* untuk prediksi data harian, yang akan diterapkan pada model *Support Vector Regression* (SVR).

B. Fungsi Kernel

Fungsi *Kernel* digunakan oleh *Support Vector Regression* (SVR) untuk memapping data dalam pola non linier ke dalam pola yang linier dengan memindahkan data ke dimensi yang lebih tinggi, di mana *hyperplane* dapat memisahkan kelas data. Fungsi *kernel* yang biasa digunakan yaitu: *Linier*,

Radial Basis Function (RBF), *Polynomial* dan *Sigmoid*. DTREG menyediakan fitur fungsi kernel tersebut yang bisa dipilih oleh user untuk menganalisa data. Penelitian ini menggunakan ke empat fungsi kernel tersebut dalam model-model yang akan dibuat, untuk kemudian dibandingkan hasil prediksinya.

C. Model *Support Vector Regression* (SVR)

Dari kombinasi *pra preparing data time series* oleh DTREG, pemilihan jenis SVR serta fungsi *kernel* yang digunakan, terdapat 16 model yang dapat dibuat.

3.6. Metode Evaluasi Hasil

Hasil eksperimen akan dievaluasi berdasarkan nilai *Error* yang dihasilkan oleh jaringan. *Error* terkecil yang diharapkan. *Result Error* yang ditampilkan oleh DTREG berupa *Root Mean Square Error* (RMSE), *Maximum Error* (ME), *Maximum Absolute Error* (MAE), *Maximum Absolute Percentage Error* (MAPE). Pada penelitian ini *Error* dievaluasi dengan penekanan pada *Root Mean Square Error* (RMSE) dalam perbandingan hasil dari model-model jaringan. Model dengan nilai *Error* RMSE terkecil sebagai algoritma dengan hasil terbaik. Metode perhitungan *Error* dapat dirumuskan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum (y_t - \hat{y}_t)^2}{T}} \quad (4.5)$$

$$MAE = \frac{\sum |y_t - \hat{y}_t|}{T} \quad (4.4)$$

$$MAPE = \frac{1}{T} \sum \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100 \quad (4.6)$$

Keterangan:

RMSE = *Root Mean Square Error*

MAE = *Maximum Absolute Error*

MAPE = *Maximum Absolute Percentage Error*

3.7. Hasil Eksperimen Dengan Variabel Prediksi 6 Lag

A. *Multilayer Perceptron* (MLP)

Dari kombinasi *pra-pemrosesan data time series*, arsitektur topologi jaringan, fungsi aktivasi pada *hidden layer* maupun *output layer*, dan metode training, terdapat 48 model yang dapat dibuat. Data training akan dilatih pada setiap model jaringan MLP tersebut, sehingga dapat membentuk *rule*, yang kemudian akan diuji oleh data validasi. Hasil peramalan dievaluasi dengan melihat nilai *Error* (RMSE) yang dihasilkan oleh setiap model. Tabel IV.4 menampilkan hasil eksperimen dari model *Multi*

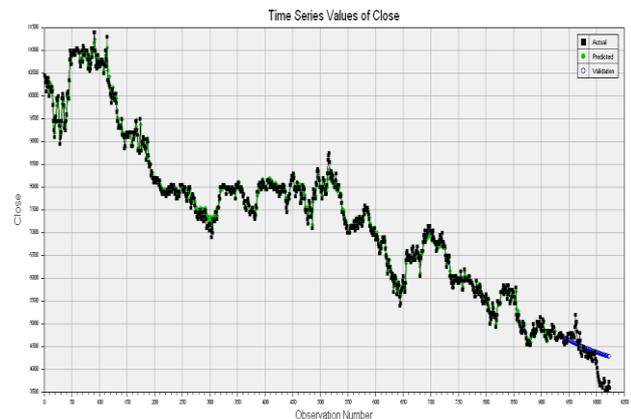
layer Perceptron MLP dengan variabel input sebanyak 6 lag.

Dari hasil eksperimen *Multilayer Perceptron* (MLP) dengan jumlah variabel prediksi sebanyak 6 lag, dapat diketahui bahwa model 28# memberikan hasil dengan nilai *Error* RMSE yang terkecil diantara 48 model yang dapat dibentuk. Spesifikasi model 28# adalah sebagai berikut :

- Model menggunakan *trendremoval* linier diikuti *stabilizevariance*.

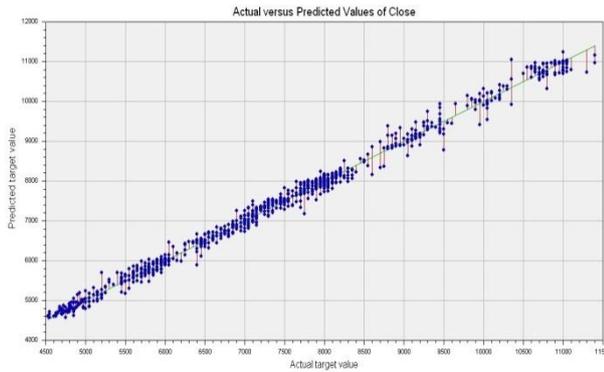
$$\begin{aligned} \text{Close_Price} &= 75874.612678 - \\ &65725.479511 * \exp(0.000084 * \text{row}) \\ \text{StabilizeVariance} &= 88.868\% \end{aligned}$$

- Model tersebut mempunyai 2 *hidden layer*, dengan topologi 6-5-4-1 yang berarti lapisan input terdiri dari 6 *neuron*, lapisan *hidden* pertama terdiri dari 5 *neuron*, lapisan *hidden* kedua dengan 4 *neuron*, dan 1 *neuron* pada lapisan *output*. Fungsi aktivasi yang digunakan pada *hidden layer* pertama dan kedua maupun *output layer* adalah fungsi *sigmoid*. Metode pelatihan yang digunakan untuk melatih data training menemukan bobot optimal untuk jaringan adalah *scaled gradientconjugate*. Model memiliki *Root Mean Square Error* (RMSE) pada hasil validasi sebesar 251,57.



Gambar 1. Grafik *Time series* Harga Saham BBCA.JK dengan MLP pada Model 28#

Pada Gambar 1 dapat dilihat grafik *time series* dari harga saham BBCA.JK pada model 28#, dimana titik-titik segi empat warna hitam adalah harga aktual saham, sedangkan titik-titik warna hijau adalah harga prediksi dari data training dengan MLP model 28#, sedangkan lingkaran biru menampilkan harga prediksi dari data validasi yang tidak digunakan untuk melatih jaringan. Sedangkan lingkaran biru menampilkan harga prediksi dari data validasi yang tidak digunakan untuk melatih jaringan.



Gambar 2. Grafik Nilai Aktual vs Nilai Prediksi Saham dengan MLP pada Model 28#

Gambar tersebut menampilkan posisi nilai prediksi dibandingkan nilai aktual dari model 28# dengan algoritma MLP. Koordinat X adalah nilai target aktual, dan Y adalah nilai prediksi. Titik-titik biru merupakan posisi pertemuan nilai aktual dan nilai prediksi. Garis diagonal merupakan prediksi sempurna di mana seharusnya nilai prediksi sama dengan nilai aktual. prediksi baik jika titik-titik biru berlokasi tidak jauh dari garis diagonal. Garis merah menunjukkan perbedaan antara nilai aktual dan prediksi, yang berupa Error atau residu.

B. Support Vector Regression (SVR) dengan Variabel Prediksi 6 Lag

Dari kombinasi pra pemrosesan data time series oleh DTREG, pemilihan jenis *Support Vector Regression* (SVR) serta fungsi *kernel* yang digunakan, terdapat 16 model yang dapat dibuat. Data training akan dilatih pada setiap model *Support Vector Regression* (SVR) tersebut, sehingga dapat membentuk *rule*, yang kemudian akan diuji oleh data validasi. Hasil peramalan dievaluasi dengan melihat nilai *Error* (RMSE) yang dihasilkan oleh setiap model.

Dari hasil eksperimen *Support Vector Regression* (SVR) dengan jumlah variabel prediksi sebanyak 6 lag yang disajikan dalam tabel IV.5, dapat diketahui bahwa model 13# memberikan hasil dengan nilai *Error* RMSE yang terkecil diantara 16 model yang dapat dibentuk. Spesifikasi model 13# adalah sebagai berikut :

- Model menggunakan *eksponentialtrendremoval* diikuti *stabilizevariance*.

$$Close_Price = 6135.495386 + 971.460917 * \exp(0.002540 * row)$$

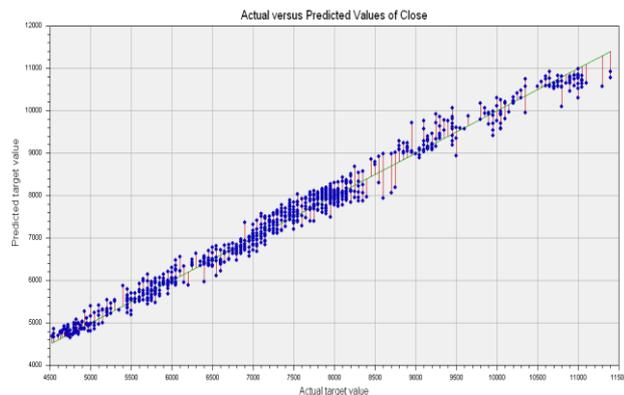
$$StabilizeVariance = 8.229\%$$
- Model menggunakan metode ϵ -SVR (*Epsilon*), dengan parameter $\epsilon = 0.00001$, $C = 1.691$, dan $\nu = 0,01$. *Kernel* menggunakan fungsi linier untuk memaping vektor *input* ke *feature space*. Jumlah *support vectoryang* digunakan oleh model sebanyak 8 buah. Model

memiliki *RootMeanSquare Error* pada hasil validasi sebesar 353,57564.



Gambar 3. Grafik Time series Harga Saham BBCA.JK dengan SVR pada Model 13#

Pada Gambar 3 dapat dilihat grafik time series dari harga saham BBCA.JK pada model 13#, dimana titik-titik segi empat warna hitam adalah harga aktual saham, sedangkan titik-titik warna hijau adalah harga prediksi dari data training dengan SVR model 13#, sedangkan lingkaran biru menampilkan harga prediksi dari data validasi yang tidak digunakan untuk melatih jaringan.



Gambar 4. Nilai Aktual dibanding Nilai Prediksi Saham dengan SVR pada Model 13 #

Gambar 4 menampilkan posisi nilai prediksi dibandingkan nilai aktual dari model 13# dengan algoritma SVR. Koordinat X adalah nilai target aktual, dan Y adalah nilai prediksi. Titik-titik biru merupakan posisi pertemuan nilai aktual dan nilai prediksi. Garis diagonal merupakan prediksi sempurna di mana seharusnya nilai prediksi sama dengan nilai aktual. prediksi baik jika titik-titik biru berlokasi tidak jauh dari garis diagonal. Garis merah menunjukkan perbedaan antara nilai aktual dan prediksi, yang berupa *Error* atau *residu*.

C. Komparasi Hasil Prediksi Variabel Input 6 Lag

Dari peramalan harga saham dengan variabel *input* 6 lag pada algoritma MLP dan SVR, diperoleh

hasil bahwa model dengan algoritma SVR lebih unggul dibandingkan peramalan dengan algoritma MLP. Nampak dalam tabel 4 hasil pengujian dengan data validasi menunjukkan nilai *Error* (RMSE) pada model dengan algoritma SVR lebih rendah dibanding model MLP.

Tabel 4. Komparasi Hasil Algoritma MLP dan SVR dengan Variabel *Input 6 Lag*

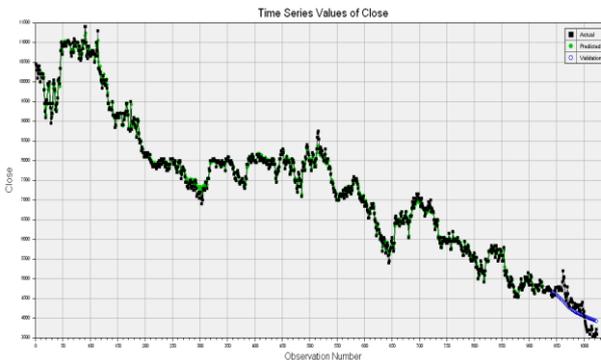
| Algoritma | Model | Data Training | | | | Data Validasi | | | |
|-----------|-------|---------------|---------|--------|------|---------------|---------|--------|------|
| | | RMSE | ME | MAE | MAPE | RMSE | ME | MAE | MAPI |
| MLP | 28# | 232,99 | 1029,25 | 168,68 | 1,85 | 471,92 | 1236,06 | 363,72 | 2,43 |
| SVR | 13# | 327,34 | 2145,28 | 235,02 | 2,76 | 353,58 | 806,67 | 298,92 | 2,06 |

3.8. Hasil Eksperimen Dengan Variabel *Input 10 Lag*

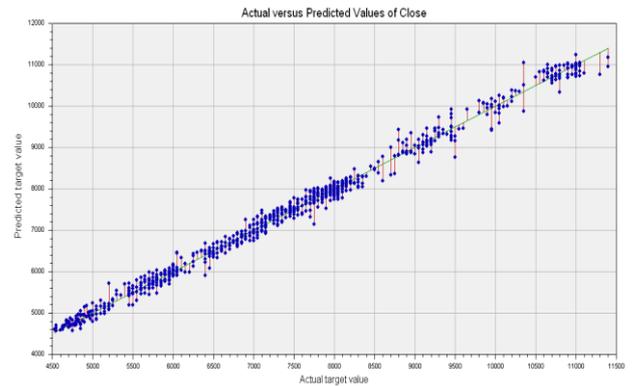
A. Multilayer Perceptron (MLP)

Dari hasil eksperimen *Multilayer Perceptron* (MLP) dengan jumlah variabel prediksi sebanyak 10 lag, dapat diketahui bahwa model 39# memberikan hasil dengan nilai *Error* RMSE yang terkecil diantara 48 model yang dapat dibentuk. Spesifikasi model 39# adalah:

- Model menggunakan *eksponensial trend removal* diikuti *stabilize variance*.
 $Close_Price = 6135.495386 + 971.460917 * \exp(0.002540 * row)$
 $StabilizeVariance = 87.229\%$
- Model tersebut mempunyai 1 *hidden layer*, dengan topologi 10-9-1 yang berarti lapisan *input* terdiri dari 10 *neuron*, lapisan *hidden* terdiri dari 9 *neuron*, dan 1 *neuron* pada lapisan *output*. Fungsi aktivasi *sigmoid* / logistik digunakan pada *hidden layer*, sedangkan *output layer* menggunakan fungsi linier. Metode pelatihan yang digunakan untuk melatih data training menemukan bobot optimal untuk jaringan adalah *traditional gradient conjugate*. Model memiliki *Root Mean Square Error* (RMSE) pada hasil validasi sebesar 497.42978



Gambar 5. Grafik *Time series* Harga Saham BBCA.JK dengan MLP pada Model 39#

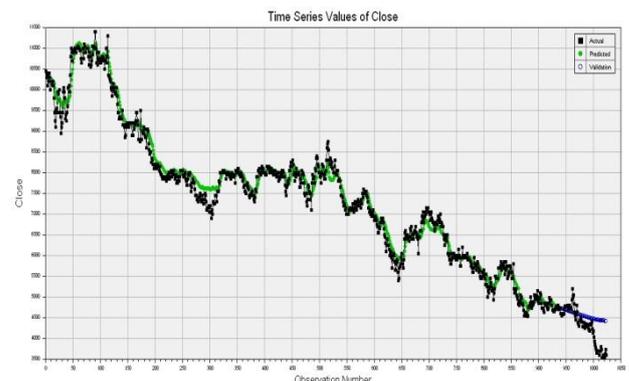


Gambar 6. Nilai Aktual dibanding Nilai Prediksi Saham dengan MLP pada Model 39#

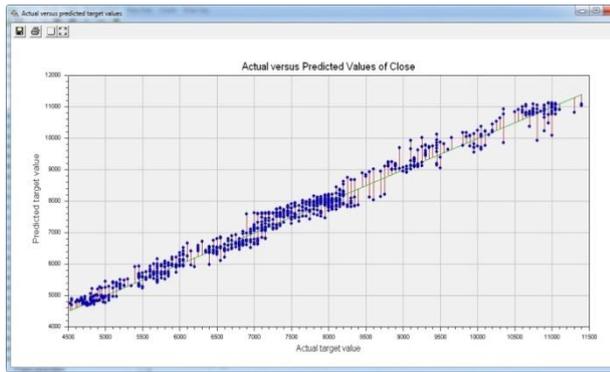
B. Support Vector Regression (SVR)

Hasil eksperimen *Support Vector Regression* (SVR) dengan jumlah variabel prediksi sebanyak 10 lag yang disajikan dalam tabel 4.8, dapat diketahui bahwa model 10# memberikan hasil dengan nilai *Error* RMSE yang terkecil diantara 16 model yang dapat dibentuk. Spesifikasi model 10# adalah sebagai berikut :

- Model menggunakan *eksponensial trendremoval* diikuti *stabilizevariance*.
 $Close_Price = 6135.495386 + 971.460917 * \exp(0.002540 * row)$
 $StabilizeVariance = 87.229\%$
- Model menggunakan metode ϵ -SVR (*Epsilon*), dengan parameter $Epsilon = 0.25$, $C = 100$, $Gamma = 4.86$ dan $P = 1,2445$. Kernel menggunakan fungsi *Radial Basis Function* (RBF) untuk memaping vektor input ke *feature space*. Jumlah *support vector* yang digunakan oleh model sebanyak 11 buah. Model memiliki *Root Mean Square Error* (RMSE) pada hasil validasi sebesar 317.88973



Gambar 7. Grafik *Time series* Harga Saham BBCA.JK dengan SVR pada Model 10#



Gambar 8. Nilai Aktual dibanding Nilai Prediksi Saham dengan SVR pada Model 10#

C. Analisa Hasil dan Komparasi Prediksi Lag 10

Dari peramalan harga saham dengan variabel prediksi 10 lag pada algoritma *Multi layer Perceptron* (MLP) dan *Support Vector Regression* (SVR), diperoleh hasil bahwa model dengan *Support Vector Regression* (SVR), lebih unggul dibandingkan peramalan dengan algoritma *Multilayer Perceptron* (MLP). Nampak dalam tabel 5 hasil pengujian dengan data validasi menunjukkan model *Support Vector Regression* (SVR) memiliki nilai *Error* (RMSE) lebih kecil dibanding model *Multi layer Perceptron* (MLP).

Tabel 5. Komparasi Hasil Algoritma MLP dan SVR dengan Variabel Input 10 Lag

| Algoritma | Model | Data Training | | | | Data Validasi | | | |
|-----------|-------|---------------|---------|--------|------|---------------|---------|--------|------|
| | | RMSE | ME | MAE | MAPE | RMSE | ME | MAE | MAPE |
| MLP | 39# | 219,0 | 2087,04 | 148,81 | 1,69 | 497,43 | 1207,01 | 411,65 | 2,82 |
| SVR | 10# | 381,75 | 2269,1 | 268,79 | 3,13 | 317,89 | 639,04 | 262,72 | 1,81 |

D. Analisa Hasil dan Komparasi Model Lag 6 dan Lag 10

Perbandingan antara penggunaan jumlah variabel prediksi dengan 6 lag dan 10 lag pada prediksi dengan algoritma *Multi layer Perceptron* (MLP) dan *Support Vector Regression* (SVR), hasilnya dapat dilihat pada table IV-10. Dalam hal keakuratan prediksi, model *Neural Network* algoritma *Support Vector Regression* (SVR) dengan variabel prediksi 10 lag lebih akurat diantara semua model lainnya. Dengan nilai *Error* RMSE yang paling kecil dibandingkan semua model, baik model SVR dengan variabel prediksi 6 lag maupun model MLP dengan 6 lag maupun 10 lag. Namun dapat diperhatikan di sini, model *Multi layer Perceptron* (MLP) dengan variabel prediksi 6 lag lebih unggul dibandingkan model dengan algoritma yang sama dengan 10 lag. Dan model *Support Vector Regression* (SVR) lebih akurat dibandingkan model *Multi layer Perceptron* (MLP) pada model 6 lag maupun 10 lag.

Tabel 6. Komparasi Hasil MLP dan SVR dengan Variabel prediksi 6 Lag dan 10

| Variabel Input | Algoritma | Model | Data Training | | | | Data Validasi | | | |
|----------------|-----------|-------|---------------|---------|--------|------|---------------|---------|--------|------|
| | | | RMSE | ME | MAE | MAPE | RMSE | ME | MAE | MAPE |
| 6 Lag | MLP | 28# | 232,99 | 1029,25 | 168,68 | 1,85 | 471,92 | 1236,06 | 363,72 | 2,43 |
| | SVR | 13# | 327,34 | 2145,28 | 235,02 | 2,76 | 353,58 | 806,67 | 298,92 | 2,06 |
| 10 Lag | MLP | 39# | 219,0 | 2087,04 | 148,81 | 1,69 | 497,43 | 1207,01 | 411,65 | 2,82 |
| | SVR | 10# | 381,75 | 2269,1 | 268,79 | 3,13 | 317,89 | 639,04 | 262,72 | 1,81 |

3.9. Implikasi Penelitian

Hasil dari penelitian ini akan membawa dampak terhadap sistem, aspek manajerial maupun terhadap penelitian selanjutnya.

A. Implikasi Terhadap Sistem

Hasil dari penelitian ini dapat diterapkan untuk meramalkan harga saham LQ45 IDX khususnya Bank Central Asia Tbk. Untuk para investor yang berminat menggunakan metode *Neural Network* untuk memprediksi harga saham, dapat menggunakan software prediksi seperti DTREG, maupun software prediksi lainnya. Tentu saja dibutuhkan perangkat keras yang memadai seperti PC atau Laptop yang mempunyai spesifikasi yang tinggi. Sebab jika terlalu rendah spesifikasi *hardware*, terutama RAM dan *processor*, seringkali tidak optimal saat *running software* prediksi. Pelatihan pengembangan model dan sosialisasi penggunaan aplikasi juga perlu dilakukan jikalau model prediksi akan diterapkan.

B. Implikasi Terhadap Manajerial

Hasil penelitian ini, memberi implikasi kepada investor yang akan mengambil keputusan baik untuk menjual atau membeli saham, terutama saham Bank Central Asia Tbk, ataupun saham perusahaan lainnya. Bahwa teknik data mining dengan *Neural Network* dapat diterapkan sebagai alternatif untuk meramalkan harga saham, di samping teknik-teknik yang lain. Jika dilihat dari hasil dalam tabel IV.10, *Maximum Error* (ME) pada model *Support Vector Regression* (SVR) hanya 639,04 dengan data validasi selama 3 bulan terakhir, menunjukkan bahwa metode peramalan cukup akurat.

Pihak manajemen sebaiknya menganggarkan dana untuk pengadaan *hardware* yang memadai dan software prediksi seperti DTREG, atau yang lainnya. Selain itu juga harus mengelola SDM yang ditugaskan untuk mengembangkan dan mengoperasikan aplikasi prediksi.

C. Implikasi Terhadap Penelitian Selanjutnya

Penelitian ini menghasilkan peramalan yang lebih baik pada model dengan variabel *input* 10 lag dibandingkan pada *input* 6 lag, terutama pada model algoritma *Support Vector Regression* (SVR) jika dibandingkan dengan model *Multi layer Perceptron* (MLP). Namun pada model *Multi layer Perceptron*

(MLP) model *input 6 lag* hasilnya lebih akurat dibandingkan *10 lag*. Hasil penelitian ini memberi dampak bagi penelitian selanjutnya untuk meneliti pengaruh jumlah variabel *input* terhadap hasil peramalan baik menggunakan *Neural Network* maupun metode pendekatan data mining yang lain. Sehingga dalam membuat model perlu dipertimbangkan dengan hati-hati berapa jumlah variabel *input* yang akan digunakan dalam prediksi.

IV. KESIMPULAN

Dalam penelitian ini dilakukan perbandingan dua model variabel prediksi 6 lag dan 10 lag pada algoritma *Multilayer Perceptron (MLP)* dan *Support Vector Machines (SVM)* untuk mendapatkan algoritma terbaik yang dapat meramalkan/memprediksi indeks harga saham. Pada model yang menggunakan algoritma MLP hasil dari setiap model sangat bergantung pada pemilihan jumlah variabel prediksi, topologi jaringan, fungsi aktivasi, metode pelatihan jaringan. Sedangkan pada algoritma SVM bergantung pada pemilihan jenis metode regresi, fungsi kernel dan parameter-parameter bebas yang ditentukan. Dan kesimpulan yang dapat diambil dari penelitian ini adalah:

- 1) Algoritma *Multilayer Perceptron (MLP)* dan algoritma *Support Vector Machines (SVM)* dapat diterapkan untuk memprediksi harga saham.
- 2) Penerapan algoritma *Support Vector Machines (SVM)* 10 lag menghasilkan keputusan yang lebih akurat dibandingkan dengan algoritma *Multilayer Perceptron (MLP)* baik di lag 6 maupun di lag 10 dengan nilai RMSE (*Root Mean Squared Error*) paling kecil 317,89, ME = 639,04, MAE = 262,72, dan MAPE=1,81 dalam meramalkan/memprediksi indeks harga saham perbankan (BBCA.JK).

DAFTAR REFERENSI

- [1] IDX. (2010) IDX Indonesia Stock Exchange Bursa Efek Indonesia. [Online]. <http://www.idx.co.id/id-id/beranda/informasi/bagiinvestor/saham.aspx>
- [2] R Abirami and M S Vijaya, "An Incremental Learning Approach for Stock Price Prediction Using Support Vector Regression," *International Journal of Research and Reviews in Artificial Intelligence (IJRAI) Vol. 1, No. 4*, pp. 81-85, December 2011.

- [3] Akbar Esfahanipour and Werya Aghamiri, "Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis," *Journal Expert Systems with Applications: An International Journal*, vol. 37, no. 7, pp. 4742-4748, July 2012.
- [4] Oded Maimon and Lior Rokach, *Data Mining and Knowledge Discovery Handbook; Second Edition*. New York : Springer , 2010.
- [5] Florin Gorunescu, *Data Mining Concept Model Technique: Intelligent Systems Reference Library, Volume 12*. Verlag Berlin Heidelberg: Springer, 2011.
- [6] Ardian Umam. (2011) <http://ardianumam.web.ugm.ac.id/>. [Online]. <http://ardianumam.web.ugm.ac.id/?p=399>
- [7] Arna Fariza, Afrida Helen, and Annisa Rasyid, "PERFORMANSI NEURO FUZZY UNTUK PERAMALAN DATA TIME SERIES," in *Seminar Nasional Aplikasi Teknologi Informasi 2007 (SNATI 2007)* , Yogyakarta, 2007, pp. 77-82.
- [8] Yulli Soelistyani, Model Peramalan Harga Saham dengan Pendekatan Neural Network Algoritma Multilayer Perceptron (MLP) dan Support Vector Regression (SVR), 2013.
- [9] Agung Wibowo and Sukrisno Mardiyanto, "Penerapan Metode Adaptive-Network-Based Fuzzy Inference System (ANFIS) Model Sugeno Untuk Memprediksi Index Saham," , 2011.
- [10] AL Muhtaram. (2011) Metris Comunity. [Online]. <http://www.metris-community.com/pengertian-data->

[mining-konsep-pdf/](#)

- [11] Carlo Verrellis, *Business Intelligence: Data Mining and Optimization for Decision Making*. Cornwall: Wiley, 2009.
- [12] S.D Balkin and J.K Ord, "Automatic neural network modeling for univariate time series," *International Journal of Forecasting*, vol. 16, pp. 509-515, 2000.
- [13] Xindong Wu and Vipin Kumar, *The Top Ten Algorithms in Data Mining*. New York: CRC Press, 2009.
- [14] Kusriani and E T Luthfi, *Algoritma Data Mining*. Yogyakarta: Andi Publishing, 2008.
- [15] Max Bramer, *Principles of Data Mining*. Verlag London: Springer, 2007.
- [16] C.R.Kothari, *Research Methodology Methods and Techniques*. India: New Age International Limited, 2004.
- [17] Rosario Capparuccia, Renato De Leone, and Emilia Marchitto, "Integrating support vector machines and neural networks," *Journal Neural Networks*, vol. 20, no. 5, pp. 590-597, July 2007.
- [18] Mauridhi Heri Purnomo, *Dasar Algoritma Cerdas*. Surabaya: PENS ITS, 2002.
- [19] Muchlisin Riadi. (2012, Desember) *Kajian Pustaka*. [Online]. <http://www.kajianpustaka.com/2012/12/pengertian-dan-jenis-jenis-saham.html#.UaZWzKLrzzw>
- [20] Thomas L Saaty, "Decision making with the analytic hierarchy process," *Int. J. Services Sciences*, vol. Vol I, No.1, pp. 83-98, 2008.
- [21] Irfan Subakti, *Sistem Pendukung Keputusan (Decision Support Sistem)*. Surabaya: Institut Teknologi Sepuluh Nopember, 2002.
- [22] Romi Satria Wahono. (2012, Aug.) <http://romisatriawahono.net>. [Online]. <http://romisatriawahono.net/2012/08/07/kiat-menyusun-kerangka-pemikiran-penelitian/>