Scientific Journal of Informatics
Vol. 5, No. 2, Nov 2018

p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Forecasting Inflation Rate Using Support Vector Regression (SVR) Based Weight Attribute Particle Swarm Optimization (WAPSO)

Erlin Mega Priliani[1], Anggyi Trisnawan Putra[2], Much Aziz Muslim[3]

[1,2,3] Computer Sceince Departement, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang, Indonesia
Email: [1]erlinmega@gmail.com, [2]anggy.trisnawan@mail.unnes.ac.id, [3]a212muslim@yahoo.com

Abstract

Data mining is the process of finding patterns or interesting information in selected data by using a particular technique or method. Utilization of data mining one of which is forecasting. Various forecasting methods have progressed along with technological developments. Support Vector Regression (SVR) is one of the forecasting methods that can be used to predict inflation. The level of accuracy of forecasting is determined by the precision of parameter selection for SVR. Determination of these parameters can be done by optimization, to obtain optimal forecasting of SVR method. The optimization technique used is Weight Attribute Particle Swarm Optimization (WAPSO). The use of WAPSO can find optimal SVR parameters, so as to improve the accuracy of forecasting. The purpose of this research is to implement SVR and SVR-WAPSO to predict the inflation rate based on Consumer Price Index (CPI) and to know the level of accuracy. The data used in this study is CPI Semarang City period January 2010-February 2018. Implementation experiments using Netbeans 8.2 gives results, SVR method has an accuracy of 94.654%. SVR-WAPSO method has an accuracy of 97.459%. Thus, the SVR-WAPSO method can increase the accuracy of 2.805% of a single SVR method for inflation rate forecasting. This research can be used as a reference for the next researcher can make improvements in determining the range of SVR parameters to get the value of each parameter more effective and efficient to get more optimal accuracy.

Keywords: Data Mining, Support Vector Regression, Particle Swarm Optimization, Inflation

1. INTRODUCTION

Data mining is a process used to find data that has not been known by the user with a model so that can be understood and used as the basis for decision making [1]. Utilization of data mining technology one of which is forecasting. Forecasting is a process to estimate some future needs that include the needs in quantity size, quality, time and location required in order to meet the demand for goods or services [2].

Time series data is a type of data that is often developed for forecasting cases. Forecasting that uses time series data in its development shows that forecasting accuracy can be improved by combining multiple models with combinations rather than using only one of the best models [3]. Frequently used forecasting methods for nonlinear time series data cases include Artificial Neural Networks (ANN), Threshold Autoregressive (TAR), Autoregressive Conditional Heteroscedastic (ARCH), and Support Vector Regression (SVR) [4].

The basis of the SVR method is to convert data into higher dimensions based on certain functions [5]. A method can be used to construct an input data mapping into a high-dimensional feature space, ie by using kernel functions. The use of the kernel function is able to detect linear relationships within the feature space with easy and efficient learning algorithms. The SVR function is divided into 2 ie linear and nonlinear. Forecasting includes nonlinear SVR functions. SVR has been investigated as an alternative technique for estimating complexity of engineering analysis [6].

SVR parameters must be considered, because the accuracy of forecasting is determined by the precision of parameter selection for SVR. Determination of these parameters can be done by optimization, to obtain optimal forecasting of an SVR. Current optimization techniques are Ant Colony Optimizer (ACO), Evolutionary Programming (EP), Differential Evolution (DE), Genetic Algorithms (AG), and Particle Swarm Optimization (PSO) [7].

Sidhartha's study [8] proves that PSOs perform best in optimization techniques compared to others. PSO is a method with a population-based approach to solve discrete and continuous optimization problems, which, when compared with other heuristic algorithms, is easier to implement and few parameters need to be determined [9]. The PSO model continues to be developed to get better optimization. The PSO model developed in this study is PSO with the addition of inertia weight attribute so called WAPSO.

WAPSO is used to reduce speed during the iteration, which allows the population to converge the target points more accurately and efficiently than the original algorithm. High inertial weight values can add to the global search portion (global exploration), while the lower value emphasizes local search. In order not to over-emphasize too high or too low a value, and keep searching for new search areas in a certain dimensional space, it is necessary to look for an inertial weight value that is equally guarded by global and local searches.

Evaluation process model is conducted to determine the performance of the method used by knowing how similar the results of forecasting and actual data. One way evaluation can be used is Mean Absolute Percentage Error (MAPE). MAPE is used to calculate the error rate for forecasting. Forecasting results are stated to have high accuracy if the MAPE value is less than 10% and is declared inaccurate if the MAPE value is more than 50% [10].

Accurate forecasting methods can be implemented in forecasting the inflation rate. Inflation is an increase in the price of goods in a continuous and widespread resulting in price increases in other goods [11]. The stability of inflation is a prerequisite for economic growth. Unstable inflation leads to uncertainty about future prices, interest rates, and exchange rates, thereby reducing economic growth [12]. The indicator used to measure the inflation rate is the Consumer Price Index (CPI).

Hidayatullah [13] uses SVR to predict consumer price index values using time series in the previous month as inputs. The Chaotic Genetic Algorithm-Simulated Annealing (CGASA) method is used to optimize SVR parameters. Forecasting results from the study has an error rate of 3.4462. Based on the above description, the purpose of this research is to implement SVR and SVR-WAPSO, and to know its accuracy in predicting inflation rate.

2. METHODS

Systematically the methods undertaken in this study starts from the stage of data processing, classification, and evaluation. Figure 1 is a flow chart of inflation rate forecasting research method using SVR and SVR-WAPSO methods, and to determine the level of accuracy.

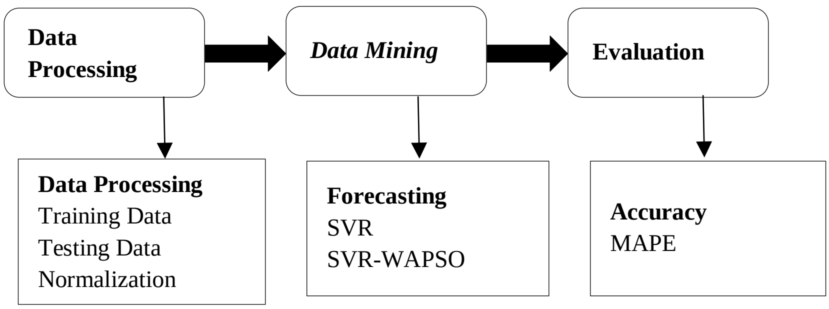

Figure 1. The flowchart of inflation rate forecasting research

2.1. Data Processing

Stages of data processing in this study using several processes such as grouping data that is divided into two data and training data test. Then the second data is done normalization process. Normalization of data is a process to scale the data so that a data is within a certain range of values. The goal is to standardize the range of features of the data, ie to make the data be in the same range. This is because the data does not necessarily have the same range, so the need for normalization. In addition, data normalization aims to reduce the level of error in computing, thus increasing the accuracy of regression. Normalization will change the data in the range of values between 0 and 1. The normalization process can be generated by the following equation.

$$\boldsymbol{x'} = \frac{x - min(x)}{max(x) - min(x)} \quad (1)$$

Information:

x' : Result of normalization

x : Current data
min (x) : Smallest data
max (x) : The largest data

After the data is normalized, at the stage before the end of the calculation using SVR-WAPSO, then there will be a process of denormalization at the time after determining the forecasting function. Denormalization is done to change the forecasting results into actual numbers. This is done because the results are still within the range of 0 to 1. The process of denormalization is done using the following equation.

$$x = x'*(\max(x) - \min(x)) + \min(x) \tag{2}$$

Information:
x' : Result of normalization
x : The value to be normalized

2.2. Forecasting

This study applies SVR and SVR-WAPSO methods to predict the inflation rate, and to know the accuracy of each method in predicting inflation.

2.2.1 Support Vector Regression (SVR) Method

The concept of the SVR method is based on minimizing the risk of error by estimating the functional sutau by minimizing the error value. The data used is data that has been normalized. Here are the steps to forecast the inflation rate using the SVR method.

1. Initialization of SVR parameters are lamda (λ), gamma (γ), complexity (C), epsilon (ε), sigma (δ). Then initialization of initial values αi * and αi each is 0 and maximum iteration.
2. Calculate the distance between data for training data and test data, with the following formula:
   $$\text{Distance} = |\text{datai-dataj}|^2 \tag{3}$$
3. Calculating the Hessian Matrix with the following equation:
   $$[R]_{ij} = K(x_i, x_j) + \lambda^2 \text{ untuk } i,j = 1, ..., n \tag{4}$$
4. The process of sequential learning. For each data, i = 1, ..., n, do with the following steps:
   a. Calculating Error value
   $$E_i = y_i - \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)\text{Rij} \tag{5}$$
   b. Calculate the value of $\delta\alpha_i^*$ and $\delta\alpha_i$
   $$\delta\alpha_i^* = \min\{\max[\gamma(E_i - \varepsilon), -\alpha_i^*], C - \alpha_i^*\} \tag{6}$$
   $$\delta\alpha_i = min\{max[\gamma(-E_i - \varepsilon), -\alpha_i], C - \alpha_i\} \tag{7}$$
   c. Calculate the value of α_i* and α_i
   $$\alpha i^* = \alpha i^* + \delta\ \alpha i^* \tag{8}$$
   $$\alpha i = \alpha i + \delta\ \alpha i \tag{9}$$
   d. Convergence has occurred, when it reaches maximum iteration or $max\ |\delta\alpha_i^*| < \varepsilon$ and $max\ |\delta\alpha_i| < \varepsilon$ then the process stops. If

you do not meet these requirements then repeat the sequential learning process in step four.

5. Forming the forecasting function. The forecasting function is used to predict the target value in the test data, with the following equation:
   $f(x) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)\ (K(x_i, \mathrm{xj}) + \lambda^2)$ (10)
6. After forming the forecasting function, it will get the output is still normalized. Then do the process of denormalization for its output back to its original value.
7. Calculate the error value using MAPE calculation.
8. Done

In this study the kernel used is the Gaussian Kernel or Radial Basis Function (RBF) because it provides the best performance to predict the load compared to other kernels [14]. Here is the equation of the RBF kernel

$K(x_{\mathrm{i}},x_{\mathrm{j}}) = exp(-\|x-x_i\|\ 2\sigma^2\ )$ (11)

Description:

$\mathrm{data}_{\mathrm{i}}$	: Data after
$\mathrm{data}_{\mathrm{j}}$	: Data before
$[\mathrm{R}]_{\mathrm{ij}}$	: Matrix *Hessian*
$K(x_i, \mathrm{xj})$	: Kernel used
λ^2	: Scalar variable
n	: Number of data
$\mathrm{E_i}$	: Error of i data
$\mathrm{y_i}$	: Actual value of i data
α	: Non-negative vector of lagrange coefisien
α_i^*	: *Lagrange multipliers*
x	: Value of data feature used for forecasting
xi	: Value of training and tasting data
σ	: Value of radial base

The flowchart of the inflation rate forecasting process using the SVR method is shown in Figure 2.

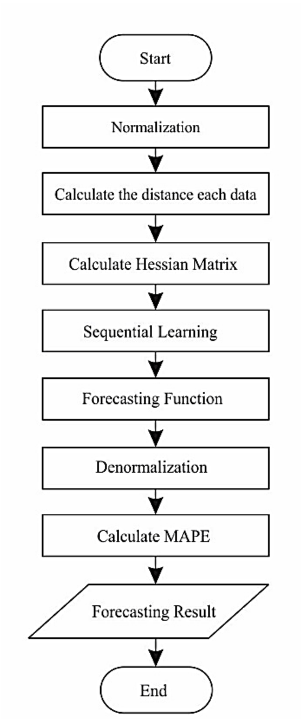

Figure 2. Flow chart of calculation of SVR method

2.2.1 SVR-WAPSO Method

SVR optimization using WAPSO aims to produce the lowest error rate in forecasting the inflation rate. WAPSO performs optimization on SVR parameters so that expected combination of SVR parameters will improve the accuracy of forecasting. Explanation of steps using SVR-WAPSO method is as follows.

1. Initialization Beginning
   The first step in the SVR-WAPSO stage is by initializing the form: maximum SVR iteration, PSO particle population number, SVR parameter range, kernel used.
2. Initialization of Early Particles
   The particle has position and velocity, for its initial value of 0. Because the particles have not moved.
3. Particle Evaluation (SVR) First
   Conducting particle evaluation using SVR method, SVR algorithm steps as in Figure 2.
4. Update ω
   $$\omega = \frac{\omega max - iter * (\omega max - \omega min)}{itermax} \quad (12)$$
5. Update Speed
   $$vid' = \omega * vid + c1 * r1 * (Pbest - xid) + c2 * r2 * (Gbest - xid) \quad (13)$$
6. Update Position

The equations of the particle position update are as follows:

$xid'=xid+vid$ (14)

7. Second Particle Evaluation (SVR)
8. Update Pbest
9. Update Gbest
10. Optimal SVR parameters
11. Forecasting using optimal SVR parameters
12. Done

Description:

ω	: Weight of inertia
Iter	: Now iteration
$c1$ dan $c2$	: Acceleration coeficient
$r1$ dan $r2$	: Random numbers selected from [0-1]
Pbest	: Best previous position
Gbest	: Best particle position in D dimension
xid’	: Newest particle position
xid	: Old particle postion (previous)
vid	: Latest particle velocity (after updating)

SVR particle evaluation is the step used to gain fitness in the SVR-WAPSO process. Fitness is a number that shows how close the forecasting compares with the actual results. The SVR particles evaluated were lamda (λ), gamma (γ), complexity (C), epsilon (ε), sigma (δ). The description of forecasting method using SVR-WAPSO method can be seen in Figure 3.

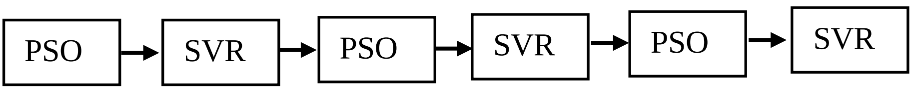

Figure 3. The description of forecasting method

2.3. Evaluation Step

Evaluation stage is useful to know the accuracy of forecasting result. Accuracy sought by way 100% minus MAPE. MAPE is formulated with the following equation.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{A_i-R_i}{A_i}\right| \quad (15)$$

Where A*i* is the actual result and R*i* is the result of forecasting.

3. RESULTS AND DISCUSSION

This study uses the data of Consumer Price Index (CPI) as much as 98 data period January 2010 to February 2018. For example, as shown in Table 1.

Table 1. Consumer price index (CPI)

Month	CPI
January 2016	122.25
February 2016	121.88
March 2016	122.35
April 2016	121.74
May 2016	121.89
June 2016	122.42
July 2016	123.70
August 2016	123.44
September 2016	123.60
November 2016	124.34
Desember 2016	124.59
January 2017	125.97
February2017	126.53
March 2017	126.35
April 2017	126.63

The data is divided into two namely the training data and test data, with a ratio of 50% each. Next is to forecast the inflation rate using SVR and SVR-WAPSO methods. The results of the implementation of the SVR method to obtain the best parameters are shown in Table 3.

Table 3. Results of the SVR best parameters

C	ε	σ	γ	λ
0.01	0.000013	0.911	0.4	0.9

Based on the best parameters of the SVR, inflation forecast in March of 113.95. The accuracy of the SVR method is shown in Table 4.

Table 4. SVR forecasting results

Month	Data Actual	Forecast	MAPE (%)	Accuracy (%)
March2018	130.94	113.95	5.346	94.654

The results of the implementation of the SVR-WAPSO method to obtain the best parameters are shown in Table 5.

Table 5. Best SVR-WAPSO parameter results

C	ε	σ	γ	λ
0.01	0.00007	0.7775	0.1	0.1

Based on the best parameters of the SVR-WAPSO, yielding inflation forecast in March of 123.95. The accuracy of the SVR-WAPSO method is shown in Table 6.

Table 6. SVR-WAPSO forecasting results

Month	Data Actual	Forecast	MAPE (%)	Accuracy (%)
March 2018	130.94	123.95	2.541	97.459

After getting the result of inflation rate forecasting using SVR and SVR-WAPSO method, the comparison of accuracy level of both methods can be seen in Table 7.

Table 7. Results of SVR comparison with SVR-WAPSO

Month	Data Actual	Forecast	
		SVR	SVR-WAPSO
March 2018	130.94	113.95	123.95
MAPE (%)		5.346	2.541
Accuracy(%)		94.654	97.459

4. CONCLUSION

The data used in the implementation of inflation rate forecasting is Consumer Price Index data obtained from the Central Bureau of Statistics from January 2010 to February 2018. Forecasting using Support Vector Regression (SVR) for March was 113.95, with the best model obtained with SVR C parameter equal to 0.01, ε of 0.000013, γ of 0.4, λ of 0.9 and σ of 0.911. While forecasting for March using SVR-WAPSO method result 123.95, and obtained the best model with SVR parameter is C equal to 0.01, ε of 0.00007, γ of 0.1, λ of 0.1, and σ of 0.775. Forecasting using SVR method has an accuracy of 94.654%, while forecasting using SVR-WAPSO method has an accuracy of 97.459%. Thus, the SVR-WAPSO method can increase the accuracy of 2.805% of the SVR method in inflation rate forecasting.

5. REFERENCES

[1] Sugiharti, E., Firmansyah, S. & Devi, F.R. (2017). Predictive Evaluation of Performance of Computer Science Students of Unnes using Data Mining Based on Naïve Bayes Classifier (NBC) Algorithm. *Journal of Theoretical and Applied Information Technology*, *95*(4), 902-911.

[2] Hani'ah, U., Arifudin, R., & Sugiharti, E. (2016). Implementasi Adaptive Neuro-Fuzzy Inference System (Anfis) untuk Peramalan Pemakaian Air di Perusahaan Daerah Air Minum Tirta Moedal Semarang. *Scientific Journal of Informatics, 3*(1), 76-87.

[3] Zaier, I., Shu, C., & Chebana, F. (2010). Estimation of Ice Thickness on Lakes Using Artificial Neural Network Ensembles. *Journal of Hidrology, 383* (3-4), 330-340.

[4] Alwee, R., Shamsuddin, S.M., & Sallehuddin, R. (2013). Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models Improved by Particle Swarm Optimization for Property Crime Rates Forecasting with Economic Indicators. *The Scientific World Journal, 2013*(2013), 1-11.

[5] Christianini, N., & Taylor, J.S. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Method.* Cambridge: Cambridge University Press.
[6] Clarke, M.S., Griebsch, J.H., & Simpson, T.W. (2005). Analysis of Support Vector Regression for Approximation of Complex Engineering Analyses. *Journal of Mechanical Design*, *127*(6), 1077-1088.
[7] Ashari, A.I, Muslim, M.A., & Alamsyah. (2016). Comparison Performance of Genetic Algorithm and Ant Colony Optimization in Course Scheduling Optimizing. *Scientific Journal of Informatics*, *3*(2), 149-158.
[8] Panda, S., & Padhy, N.P. (2007). Comparison of Particle Swarm Optimization and Genetic Algorithm for TCSC-based Controller Design. *International Journal of of Electrical, Computer, Energetic, Electronic and Communication Engineering*, *1*(3), 543-551.
[9] Santosa, Budi, & Willy, Paul. (2011). *Metoda Metaheuristik, Konsep dan Implementasi.* Surabaya: Graha Ilmu.
[10] Lewis, C.D. (1982). *Industrial And Business Forecasting Methods, A Practical Guide To Exponential Smoothing And Curve Fitting*. London: Butterworths.
[11] Bank Indonesia. (2014). *Statistik Moneter.* Jakarta: Bank Indonesia.
[12] Jha, R., & Dang, T. N. (2011). Inflation Variability and The Relationship Between Inflation and Growth. *ASARC Working Paper* 2011/08, 1-18.
[13] Hidayatullah, Sholihin, M.M, Cholissodin, I., & Perdana, R.S. (2018). Peramalan Kenaikan Indeks Harga Konsumen/Inflasi Kota Malang menggunakan Metode Support Vector Regression (SVR) dengan Chaotic Genetic Algorithm-Simulated Annealing (CGASA). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *2*(2), 857-868.
[14] Hong, W.C. (2009). Electric Load Forecasting by Support Vector Model. *Aplied Mathematical Modelling, 33*(5), 2444-2454.