SVR OPTIMIZATION WITH PSO FOR CRYPTOCURRENCY PRICE PREDICTIONS

Arifin Hidayat[1], Andi Farmadi[2], Mohammad Reza Faisal[3], Dodon Turianto Nugrahadi[4], Rudy Herteno[5]
[12345]Computer Science FMIPA ULM
A. Yani St. KM 36 Banjarbaru, South Kalimantan
Email: j1f115204@mhs.ulm.ac.id[1], andifarmadi@ulm.ac.id[2], reza.faisal@ulm.ac.id[3], dodonturianto@ulm.ac.id[4], rudy.herteno@ulm.ac.id[5].

Abstract

Cryptocurrency is the nickname given to a system that uses Cryptography technology to securely transmit data and process digital currency exchanges in a dispersed manner. A Cryptocurrency is a form of risky investment, Cryptocurrency prices are very volatile (changing) making Cryptocurrency prices need to be predicted to make a profit. Support Vector Regression (SVR) is one method for predicting time series data such as Cryptocurrency prices. However, the SVR parameters need to be optimized to get accurate results. The Particle Swarm Optimization (PSO) algorithm is implemented to determine the effect on the optimization of SVR parameters. The implementation of SVR and SVR-PSO is carried out on Bitcoin and Shiba Inu Coin Cryptocurrency data. The result of this research is that the SVR algorithm has an accuracy of 13.19082% (Bitcoin) and 68.3221% (Shiba Inu Coin). The SVR-PSO algorithm obtained an accuracy of 96.92359% (BTC) and 94.74245% (SHIB).

Keywords: *Cryptocurrency, Prediction, Time Series, Support Vector Regression, Particle Swarm Optimization*

1. INTRODUCTION

Cryptocurrency is the nickname given to a system that uses cryptographic technology to securely transmit data and process digital currency exchanges widely [1]. A Cryptocurrency is a form of investment that is risky but can provide great benefits if it can be managed properly. The value of Cryptocurrency prices is very volatile and is purely influenced by supply and demand, in contrast to ordinary currencies whose exchange rates are strongly influenced by the central bank, the exchange rate of Cryptocurrency is determined by the Cryptocurrency owners themselves. This price uncertainty is more extreme than the uncertainty of stock prices and foreign currencies because of the nature of Cryptocurrency itself. This uncertainty is certainly not liked by traders and investors because it poses a risk to their investment while this uncertainty is also unavoidable in investing. To deal with uncertainty, a prediction technique (forecasting) can be used to predict the value of the increase or decrease in Cryptocurrency prices. Forecasting is used to predict something in the future based on past data so that users can make the right decisions based on forecasting data [2].

One method that can forecast time series data is Support Vector Regression (SVR). Support Vector Regression (SVR) is a regression model of the Support Vector Machine (SVM) algorithm. Like the research conducted by [3] to predict stock prices, SVR is a machine learning algorithm that has good performance for recognizing patterns from time series data and can provide good predictive results if the values of the important parameters can be determined properly as well.

SVR parameters are usually chosen based on experience and grid search (GS), experience requires a deep understanding of SVR theory, while the computations for the GS method are large so both methods cannot guarantee an optimal solution [4]. Therefore, an optimization method is needed to determine the SVR parameters so that SVR can be optimally applied in forecasting Cryptocurrency prices. Optimization of SVR parameters has been carried out, such as research on the prediction of car sales volume with SVR-PSO [5] which states that PSO-SVR is better than GA-SVR in terms of efficiency and prediction accuracy. PSO is a method with a population-based stochastic approach to solving discrete and continuous optimization problems, which when compared to the Genetic Algorithm (GA) and other heuristic algorithms, is easier to implement and fewer parameters need to be specified [6]. Another research that is about forecasting short-term solar radiation using SVR gives the result that the parameter optimized SVR using PSO gives better forecasting results compared to ordinary SVR and SVR-GA [7].

Therefore, the goal to be achieved in this research is to find out how accure results are produced by the Support Vector Regression method that has been optimized with the Particle Swarm Optimization algorithm in predicting Cryptocurrency prices. Data used in this research are Bitcoin data and Shiba Inu coin data from investing.com.

2. RESEARCH METHODOLOGY

The flow of this research can be seen in Figure 1 below:

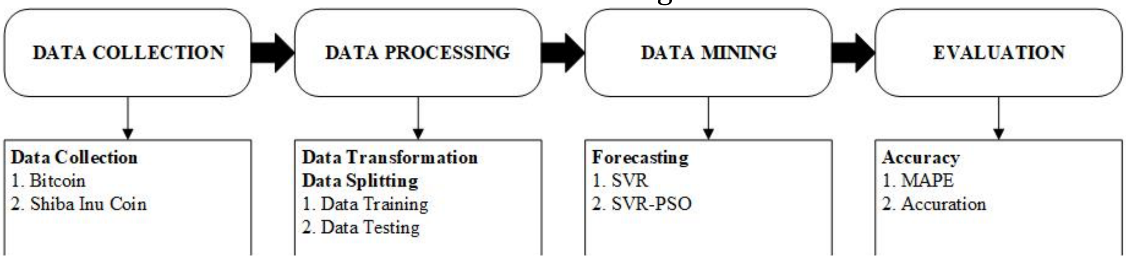

Figure 1. Research flow of stock price predictions

2.1 Data Collection

Data collection is a process to obtain Cryptocurrency data that will be used in predictions. The data is historical data from the price of Bitcoin and Shiba Inu Coin. Range of data taken from May 2021 to May 2022. The data is in the form of daily data and obtained from investing.com

2.2 Data Processing

Data processing is divided into two stages, the first is data transformation and the second is data splitting. Data transformation is to transform the data form into many features or time-series data with a certain window size. Data splitting is a process of splitting the data into training data and testing data based on a predetermined ratio.

2.3 Data Mining

Data mining phase implements forecasting using 2 algorithms namely SVR and SVR-PSO. The first algorithm is SVR which is Support Vector Regression algorithm without parameter optimization. The second algorithm is SVR-PSO which is Support Vector Regression algorithm with parameter optimization using Particle Swarm Optimization algorithm.

2.3.1 SVR Implementation

Time series data can be predicted using Support Vector Regression (SVR) algorithm. Sequential learning in SVR can be used to solve non-linear problems [5]. Before calculating with SVR algorithm, normalize the data first using this equation (1):

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

x' : result of normalization

x : current data value

Min (x) : smallest data value

Max(x) : the largest data value

Then implement SVR algorithm steps as follows:

a. Initialize SVR parameters such as complexity (C), lamda (λ), epsilon (ε), sigma (σ) and gamma (γ). Intialization of initial values and each is 0 and maximum iteration.$\alpha_i^* \alpha_i$

b. Calculate the distance between data for training data and testing data, with the following formula:

   Distance = | datai - dataj | 2 (2)

c. Calculating the Hessian Matrix with the following equation:

   $R_{ij} = K(x_i, x_j) + \lambda^2$ for i, j = 0, …, n (3)

d. The process of sequential learning. For each data, i = 0, ..., n, do with the following steps:

   1. Calculating Error value

      $E_i = y_i - \Sigma_{i=1}^{n}(\alpha_i^* - \alpha_i)R_{ij}$ (4)

   2. Calculate the value of and$\delta\alpha_i^*$ $\delta\alpha_i$

      $\delta\alpha_i^* = min\{max[\gamma(E_i - \varepsilon), -\alpha_i^*], C - \alpha_i^*\}$ (5)

      $\delta\alpha_i = min\{max[\gamma(-E_i - \varepsilon), -\alpha_i], C - \alpha_i\}$ (6)

   3. Calculate the value of andα_i^* α_i

      $\alpha_i^* = \alpha_i^* + \delta\alpha_i^*$ (7)

      $\alpha_i = \alpha_i + \delta\alpha_i$ (8)

   4. Convergence has occurred, when it reaches maximum iteration or max < ε and max < ε then the process stops. If you do not meet these requirements then repeat the sequential learning process in step four.$|\delta\alpha_i^*||\delta\alpha_i|$

e. Forming the forecasting function. The forecasting function is used to predict the target value in the test data, with the following equation:

   $f(x) = \Sigma_{i=0}^{n}(\alpha_i^* - \alpha_i)\, K(x_i, x_j) + \lambda^2$ (9)

f. 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.

g. Calculate the error value using Mean Absolute Percentage Error (MAPE).

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{A_i - R_i}{A_i} \right| \quad (10)$$

h. Done.

In this study used Radial Basis Function (RBF) kernel because compared to other kernels, this kernel provides the best performance to predict the load. Here is the equation of the RBF kernel [2]:

$$K(x_i, x_j) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (11)$$

Description:

datai : data before
dates : data after
R_{ij} : matrix hessian
$K(x_i, x_j)$: kernel used
λ : scalar variable
n : number of data
No. : error of i data
do : actual value of i data
α_i^* : non-negative vector of lagrange coefisien
α_i : lagrange multipliers
A_i : the actual result
R_i : the forecasting result
x : the value of data feature used for forecasting
x_i : the value of training and testing data
σ : the value of radial base

The Flowchart of stock price predictions using the SVR algorithm is shown in figure 2.

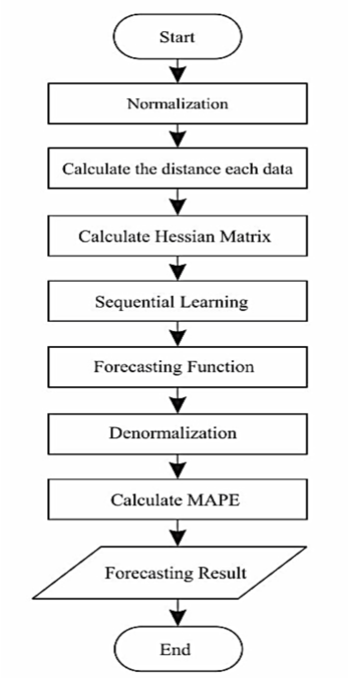

Figure 2. Flowchart of SVR algorithm

2.3.2 SVR-PSO Implementation

Forecasting using SVR-PSO will use SVR as the prediction/forecasting algorithm and PSO as optimization algorithm. The steps for the SVR-PSO algorithm are as follows [9]:

a. The first step is initialization initial such as SVR parameter range (complexity (C), lamda (λ), epsilon (ε), sigma (σ) and gamma (γ)), maximum PSO iteration, PSO particle population number, and kernel.
b. Particle initialization, the initial velocity of each particle is 0 and the initial position of each particle is random using the following equation:
   $$x = xmin + rand[0,1] * (xmax - xmin) \quad (12)$$
   x : position particle or value of SVR parameter
   $xmin$: minimum value of SVR parameter range
   $xmax$: maximum value of SVR parameter range
c. Use the SVR algorithm to evaluate the early particle (first evaluation), SVR algorithm steps as figure 2 and calculate fitness value with the following equation:
   $$fitness = \frac{1}{1+MAPE} \quad (13)$$
d. Update c1 and c2
   $$c_1 = (c_{1f} - c_{1i}) * \frac{iter}{itermax} + c_{1i} \quad (14)$$
   $$c_2 = (c_{2f} - c_{2i}) * \frac{iter}{itermax} + c_{2i} \quad (15)$$
e. Update velocity
   $$v'_{id} = v_{id} + c_1 * r_1 * (p_{id} - x_{id}) + c_2 * r_2 * (p_{gd} - x_{id}) \quad (16)$$
   Velocity clamping with rule:
   If > then v'_{id} $v_{max} v'_{id} = v_{max}$
   If < then $v'_{id} v_{min}$ $v'_{id} = v_{min}$
f. Update position
   $$x'_{id} = x_{id} + v'_{id} \quad (17)$$
   Normalizes the position (position clamping) with the rule:
   If > thenx'_{id} $x_{max} x'_{id} = x_{max}$
   If < then$x'_{id} x_{min}$ $x'_{id} = x_{min}$
g. Second particle evaluation using SVR and calculate fitness value.
h. Update Pbest.
i. Update Gbest.
j. Repeat from step 4 to step 9 until iteration reaches the maximum value.
k. Optimal SVR parameters were obtained.
l. Forecasting using Optimal SVR parameters.
m. Done.

Information:
Maximum and minimum velocity can calculate with equation:

v_{max} = k (−) $k \in [0,1] x_{max}$ x_{min} (18)

v_{min} = $-v_{max}$ (19)

Description:
$c_1$and c_2 : Acceleration coeficient
$iter$: now iteration
$itermax$: maximum iteration
$r_1$andr_2 : random number selected from [0-1]
$Pbest/p_{id}$: best previous position (local)
$Gbest/p_{gd}$: best particle position in D dimension (global)
x'_{id} : newest particle position (after update)
x_{id} :previous particle position (before update)
v'_{id} :newest particle velocity (after update)
v_{id} : previous particle velocity (before update)
The Flowchart of forecasting using the SVR-PSO algorithm is shown in figure 3.

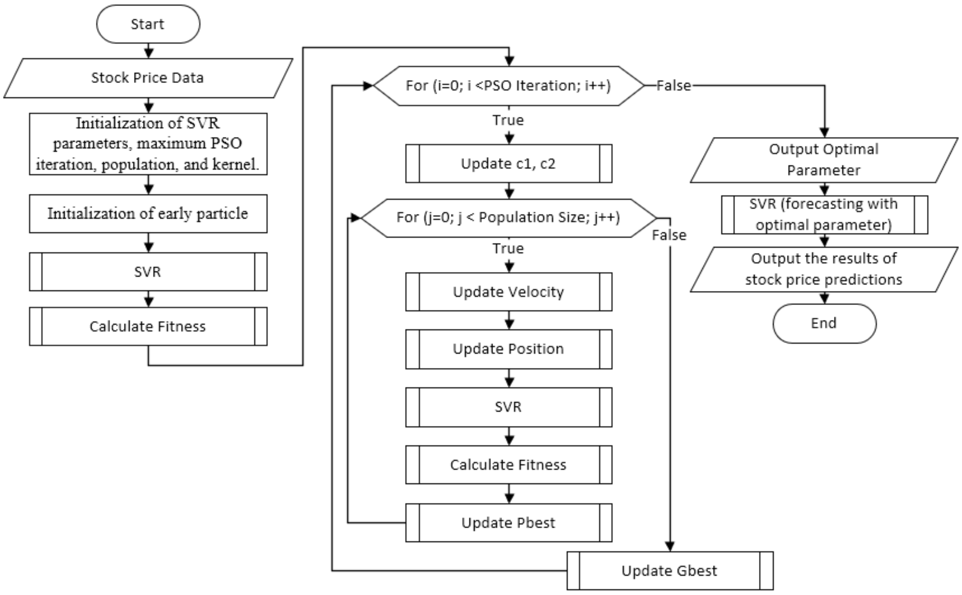

Figure 3. Flowchart of SVR-PSO algorithm

2.4 Evaluation

The evaluation will be carried out using MAPE calculations to measure the error rate and calculate the accuracy value. Accuracy is obtained from the calculation of 100% minus the MAPE value. Then compare the MAPE results and the accuracy of each method used. Then it can be determined which method is better. If the MAPE value is less than 10%, it can be said that the results are accurate[8].

3. RESULTS AND DISCUSSION

3.1 Results

The data used in this study is Bitcoin and Shiba Inu Coin price data from May 2021 to May 2022 obtained from Investing.com. The data used are Bitcoin (BTC) and Shiba Inu Coin (SHIB) data. BTC and SHIB data will be predicted using SVR and SVR-PSO algorithms.

Table 1 Amount of data collected

No	Data	Range Data	Total Data
1.	BTC	15 Mei 2021 - 18 Mei 2022	369
2.	SHIB	15 Mei 2021 - 18 Mei 2022	369
		Total	738

Data transformation is performed in data processing, where data from close price column used to make 3 feature or 3 windows size time series data become column Day-3, column Day-2, column Days-1 and column Actual. Total data will be 366 rows. Then split the data into training data and testing data with ratio 90%:10%. Total training data is 329 data and testing data is 37 data.

Table2 Bitcoin training data

Index	Date	Day-3	Day-2	Day-1	Actual
0	5/18/2021	46709	46426	43541	42897
1	5/19/2021	46426	43541	42897	36721
2	5/20/2021	43541	42897	36721	40717
...	...	...	...	...	...
328	4/11/2022	42275	42767	42138	39497

Table 3 Bitcoin testing data

Index	Date	Day-3	Day-2	Day-1	Actual
329	4/12/2022	42767	42138	39497	40078
330	4/13/2022	42138	39497	40078	41133
331	4/14/2022	39497	40078	41133	39936
...	...	...	...	...	...
365	5/18/2022	31309	29849	30438	29983

3.1.1 Calculation of SVR Algorithm

Normalize the data first with equation (1) and the result will be as bellow:

Table 4 Bitcoin training data after normalization

Index	Date	Day-3	Day-2	Day-1	Actual
0	5/18/2021	0.459976	0.452635	0.377801	0.361097
1	5/19/2021	0.452635	0.377801	0.361097	0.200897
2	5/20/2021	0.377801	0.361097	0.200897	0.304550
...	...	...	...	...	...
328	4/11/2022	0.344963	0.357725	0.341409	0.272904

Table 5 Bitcoin testing data after normalization

Index	Date	Day-3	Day-2	Day-1	Actual
329	4/12/2022	0.357725	0.341409	0.272904	0.287975
330	4/13/2022	0.341409	0.272904	0.287975	0.315340
331	4/14/2022	0.272904	0.287975	0.315340	0.284291

...	...	...	...	...	...
365	5/18/2022	0.060516	0.022645	0.037923	0.026121

Then the SVR calculation steps is initialize SVR parameters such as complexity (C) with range 100 - 500, lamda (λ) with range 0.001 - 0.1, epsilon (ε) with range 0.0001 - 0.01, sigma (σ) with range 0.001 - 2 and gamma (γ) with range 0.00001 - 0.001, initial values = 0 and = 0 and maximum iteration = 1000.$\alpha_i^*\alpha_i$

Table 6 Initial SVR parameters on Bitcoin data

C	λ	e	Σ	γ
431.08	0.061073	0.008542	0.648476	0.000982

Calculate the data with the initial SVR parameters from the SVR algorithm step 2 to step 6 and the result is:

Table 7 SVR prediction results on Bitcoin data

Index	Date	Actual	Prediction Results
329	4/12/2022	40078	32833.21
330	4/13/2022	41133	32747.17
331	4/14/2022	39936	32719.02
...	...	...	...
365	5/18/2022	29983	31339.1

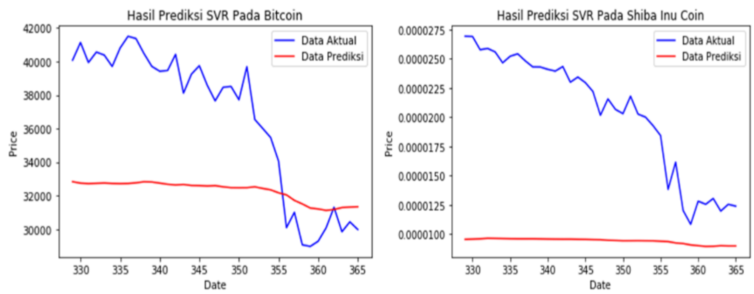

Figure 5. Graph of SVR prediction results on Bitcoin and Shiba Inu Coin

Next step is to calculate MAPE with equation (10) and the result is:

Table 8 Parameters and MAPE of SVR prediction results on stock price data

No	Data	C	λ	e	σ	γ	MAPE (%)
1.	BTC	431.08	0.061073	0.008542	0.648476	0.000982	13.19082
2.	SHIB	110.28	0.0278884	0.00895753	0.738031	0.000839818	50.3714

3.1.2 Calculation of SVR-PSO Algorithm

The SVR-PSO algorithm can be calculated with first step is initialize SVR parameters range (complexity (C) = 100 - 500, lamda (λ) = 0.001 - 0.1, epsilon (ε) = 0.0001 - 0.01, sigma (σ) = 0.001 - 2 and gamma (γ) = 0.00001 - 0.001), maximum PSO iteration = 40, PSO particle population = 30, and kernel using RBF (Radial Basis Function). Then initialize the early particles with equation (12) and set initial velocity = 0. Calculate using equation (18) and using equation (19). Evaluate the early particle with the SVR algorithm steps as shown in figure 2 and calculate the fitness value using equation (13). The result is:$v_{max} v_{min}$

Table 9 Results of first evaluation using SVR algorithm on Bitcoin data

Particle	C	λ	e	σ	γ	MAPE (%)	Fitness
1	102.92	0.05491	0.00042	1.23998	0.00058	10.2724	0.08871
2	417.08	0.05380	0.00372	1.45967	0.00016	17.9216	0.05285
3	195.68	0.02953	0.00076	1.18741	0.00079	10.8112	0.08467
...	...	...	...	...	...	...	...
30	180.44	0.06436	0.00423	0.76462	0.00028	17.1709	0.05503

Set first evaluation result as initial Pbest and set initial Gbest with a particle that has the best fitness value. The next step is to enter the PSO iteration with an update and using equations (14) and (15), update velocity with equation (16), update position using equation (17) and evaluate again with SVR algorithm. Update Pbest and Gbest after get the evaluation result. Repeat these steps until the maximum PSO iteration is complete. Then the optimal SVR parameters are obtained as below:$c_1 c_2$

Table 10 Optimal SVR parameters using SVR-PSO

No	Data	C	λ	e	σ	γ	MAPE (%)	Fitness
1.	BTC	204.21	0.06074	0.0001	1.183161	0.000504	3.076405	0.245314
2.	SHIB	322.33	0.029672	0.0001327	0.538843	0.000863	5.25755	0.1598

The last step is forecasting using optimal SVR parameters and the result is:

Table 11 SVR-PSO prediction results on Bitcoin

Index	Date	Actual	Prediction Results
329	4/12/2022	40078	41283.009
330	4/13/2022	41133	40401.374
331	4/14/2022	39936	40191.241
...	...	...	...
365	5/18/2022	29983	30942.742

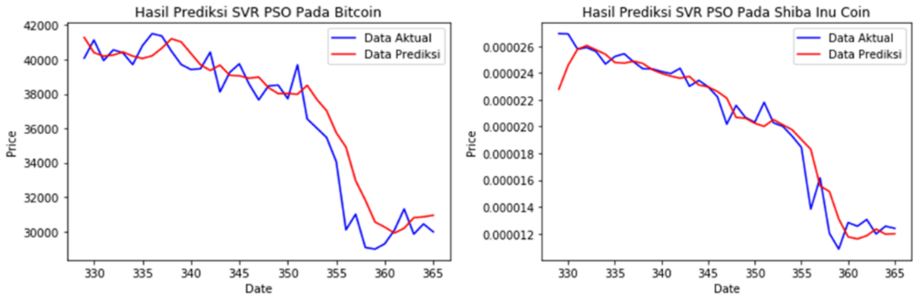

Figure 6. Graph of SVR-PSO prediction results on Bitcoin and Shiba Inu Coin

3.1.3 Calculation of Accuracy

Accuracy can calculate with equation 100% minus MAPE and the result is:

Table 12 Accuracy results

No	Algorithm	Data	MAPE (%)	Accuracy (%)
1.	SVR	BTC	13.19082	86.80918
		SHIB	50.37144	49.62856
2.	SVR-PSO	BTC	3.076405	96.92359
		SHIB	5.25755	94.74245

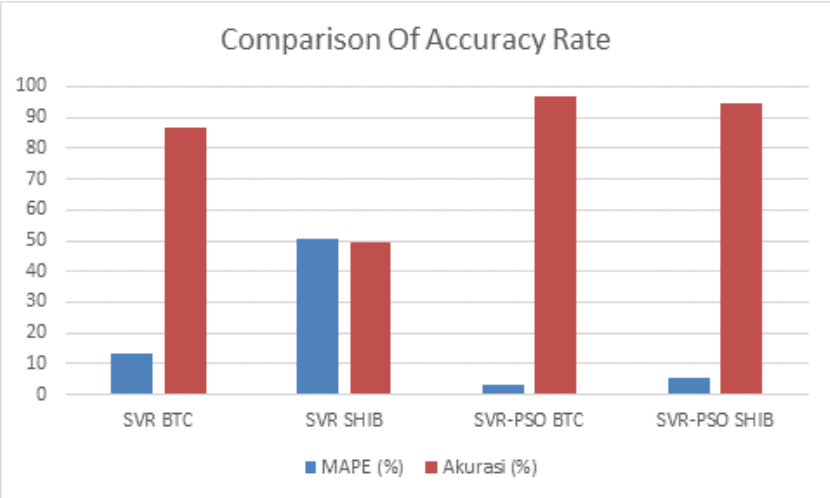

Figure 8. Graph of accuracy rate comparison

3.2 Discussion

The data used in this research is Bitcoin (BTC) and Shiba Inu Coin (SHIB) price data with each data amounting to 369 data. Each data will be implemented by 2 algorithms, namely SVR and SVR-PSO. PSO is used for parameter optimization on SVR to get maximum accuracy. MAPE obtained from the SVR algorithm on BTC and SHIB data is more than 10%. Based on the previous theory that if the MAPE is less than 10%, it can be said that the results are accurate. This shows that the results of the implementation of the SVR algorithm are not accurate. After optimizing the parameters using PSO, the accuracy results are above 90% and MAPE is below 10%. From the results of this study, it was found that the SVR that had been optimized using PSO showed better results, both in terms of the accuracy value and the average fitness value.

4. CONCLUSION

The accuracy of the Support Vector Regression (SVR) algorithm on BTC and SHIB data without parameter optimization is 86.80918% and 49.62856%, respectively. While the accuracy of the Support Vector Regression (SVR) algorithm on Bitcoin and Shiba Inu Coin with parameter optimization using the Particle Swarm Optimization (PSO) algorithm is 96.92359% and 94.74245%, respectively.

REFERENCES

[1] Shabrina Puspasari. 2020. “Perlindungan Hukum bagi Investor pada Transaksi Aset Kripto dalam Bursa Berjangka Komoditi.” Jurist-Diction, 3(1), 303-330.
[2] Kurniawan, Bayu Aji. 2018. “Peramalan Harga Bitcoin Menggunakan Backpropagation Neural Network.” Intitut Teknologi Sepuluh November.
[3] Meesad, Phayung and Rasel, Risul Islam. 2013. “Predicting Stock Market Price Using Support Vector Regression”. 2013 International Conference on Informatics, Electronics and Vision (ICIEV), pp. 1-6.
[4] Pei-you, C. & Lu, L., 2013. “Study on Coal Logistics Demand Forecast Based on PSOSVR”. 2013 10th International Conference on Service Systems and Service Management, pp. 130-133.
[5] Lu, X. & Geng, X., 2011. “Car Sales Volume Prediction Based on Particle Swarm Optimization Algorithm and Support Vector Regression”. Fourth International Conference on Intelligent Computation Technology and Automation, pp. 72-74.
[6] Hu, W., Yan, L., Liu, K. & Wang, H., 2015. “PSO-SVR: A Hybrid Short-term Traffic Flow Forecasting Method”. 2015 IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS), pp. 553- 561.
[7] Sreekumar, S., Sharma, K. C. & Bhakar, R., 2016. “Optimized Support Vector Regression Models for Short Term Solar Radiation in Smart Environment”. 2016 IEE Region 10 Conference (TENCON), pp. 1929-1932.
[8] Rusmalawati, Vera et al. 2018. "Peramalan Harga Saham Menggunakan Metode Support Vector Regression (SVR) dengan Particle Swarm Optimization (PSO)." Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer 2(5): 1980-1990.
[9] Hidayat, Arifin. 2022. "Optimasi Support Vector Regression (SVR) Dengan Particle Swarm Optimization (PSO) Untuk Peramalan Cryptocurrency."