Bulletin of Applied Mathematics and Mathematics Education
Volume 3, Number 2, 2023 pp. 55-62
https://dx.doi.org/10.12928/bamme.v3i2.9561
e-ISSN 2776-1029, p-ISSN 2776-1002

Support Vector Regression optimization with Particle Swam Optimization algorithm for predicting the gold prices

Novi Selviani, Joko Purwadi*

Mathematics Department, Universitas Ahmad Dahlan, Jl. Jend. Ahmad Yani, Tamanan, Banguntapan, Bantul, DIY 55711, Indonesia

*Corresponding E-mail: joko@math.uad.ac.id

ARTICLE INFO

Article History
Received 14 December 2023
Revised 29 December 2023
Accepted 29 December 2023

Keywords
Particle Swam Optimization
Prediction
Support Vector Regression

How to cite this article:
Selviani, N., & Purwadi, J. (2023). Support Vector Regression optimization with Particle Swam Optimization algorithm for predicting the gold prices. *Bulletin of Applied Mathematics and Mathematics Education*, *3*(2), 55-62.

ABSTRACT

This paper discusses about how to predict the gold prices from 1 January 2021 to 31 January 2023. The method used in this study is the Support Vector Regression (SVR) technique, method that was developed from the support vector machine which is used as regression approach to predict future event. From the past study already know that SVR had limitation in achieving good performance because of its sensitivity to parameters. To overcome the SVR performance problems, an optimization algorithm is proposed in this study. The PSO algorithm is applied in this study to optimize the parameters of the SVR method. The results showed that the prediction of the SVR model obtained an MSE value of 0.0035744. While in the SVR model with the PSO algorithm, the MSE value is 0.0033058.

Introduction

Investment is the act of directly or indirectly investing capital with the belief that the owner of the capital would profit from the outcomes of his investment in the future. In general, investment is defined as the expenditure or use of time, money, or energy to generate future returns. Investing is purchasing something with the expectation of reselling it at a greater price in the future. When compared to other sorts of investment products, gold is considered a simple type of investing. Gold investing is believed to be simple since gold does not need to be owned by someone with a huge income or a specific position (Setyowibowo dkk, 2022). Investors are particularly interested in investing in gold, and they do so as a long-term investment. The daily volatility of gold prices is a regular risk in gold investing (Ali, R., dkk 2020). Given the large level of investor interest, more accurate estimation or forecast of future prices is required (Syahri & Robiyanto, 2020). Prediction is a methodical prediction of what will happen in the future based on previous and present data. One of the strategies used in data mining techniques for prediction issues is the Support Vector Regression (SVR) method.

Some studies, such as those carried out by Yupie Kusumawati et al, 2022, used the Support Vector Regression (SVR) approach to forecast gold prices. In this study, the choice of features, kernels, and parameters has a significant impact on the prediction outcomes. Then, in their study, Nazish Ashfaq et al compared several machine learning regressors such as the Support Vector Regression (SVR) technique, Elastic Net, Lasso, Ridge Regression, Decision Tree, Ransac, and Extra Tree. When compared to the Decision Tree, Extra Tree, and Ransac techniques, the SVR, Lasso, Elastic Net, and Ridge Regression methods have higher accuracy and are better at predicting the stock market. Furthermore, Tran Thanh Ngoc et al, 2021 used Support Vector Machines in their investigation.

Based on previous research, the researchers intend to use the Support Vector Regression (SVR) method in conjunction with the PSO (Particle Swarm Optimization) algorithm to optimize the parameter values in the SVR method, namely the C parameters, epsilon (ε), and gamma (γ), in predicting gold prices and providing the best prediction models and results.

Method

Support Vector Machine (SVM)

SVM was first introduced by Vapnik in 1992 as one of the machine learning methods that works according to the SRM (Structural Risk Minimization) principle [15]. Support Vector Machine (SVM) is a classification method that works by finding the hyperplane with the largest margin. A hyperplane is a line separating data between classes to distribute data so that classification and regression analysis can be carried out. Margin is the distance between the hyperplane and the closest data in each class. The data closest to the hyperplane of each class is called the support vector. Illustration of SVM classification by finding a hyperplane that separates the two classes.

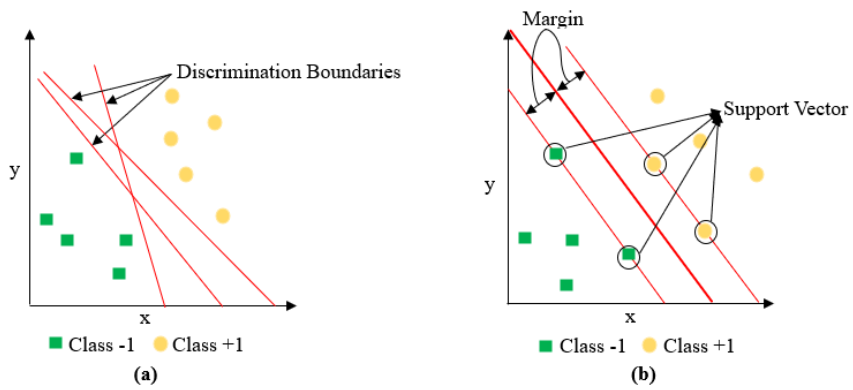

Figure 1. Illustration of SVM Classification

Figure 1(a) shows patterns belonging to two classes: +1 and -1. The -1 category patterns are denoted in green (squares) and the +1 category patterns are denoted in yellow (circles). The classification problem can be interpreted by trying to find the hyperplane that separates the two groups. Alternative discrimination boundaries can be seen in Figure 1(a) The hyperplane that best discriminates between the two classes is found by measuring the margin of the hyperplane and finding its maximum point. Figure 1(b) shows the best hyperplane, which is exactly in the middle of the two classes, while the green and yellow dots inside the black circle are the support vectors.

Suppose in a dataset there is data given by variable $\boldsymbol{x_i}$ with $\mathrm{i} = (1,2,3, \ldots, \mathrm{i})$ and there is a dataset class given by variable y_i with $\mathrm{i} = (-1, +1)$. The dataset is then divided into 2 classes by the SVM method. The first class is separated by a hyperplane with a value of +1, while the second class is separated by a hyperplane with a value of -1. The generalized hyperplane that separates the two classes can be denoted as follows.

 10.12928/bamme.v3i2.9561

$\boldsymbol{w} \cdot \boldsymbol{x_i} + b = 0$ (1)

From equation (1), the equation of the 2 hyperplane that separates the first class, and the second class is obtained, namely:

i. For the first class with value +1:

$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b = +1$$

ii. For the second class with value -1:

$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b = -1$$

with $\boldsymbol{w}$ denote the support vector weight value that perpendicular tits hyperplane, $\boldsymbol{x_i}$ the i-th data and b as the bias value.

Support Vector Regression (SVR)

Support Vector Regression (SVR) aims to find a function $f(x)$ as the hyperplane of the regression function shape that fits all input data errors and make it as thin as possible. Suppose there are n training data with $(x_i,\ y_i), i = 1, 2, 3, \dots, n$ where x_i is the input vector $x = \{x_1\ ,\ x_2, \dots, x_n\} \subseteq \mathbb{R}^n$ and the output scalar $y = \{y_1, \dots, y_n\} \subseteq\ \mathbb{R}$ and n is the number of training data. Using SVR, a function $f(x)$ will be determined that has the largest deviation ε from the actual target y_i, for all training data. If the value of ε is equal to 0 then a perfect regression equation is obtained. So the hyperplane function can be written as follows

$$f(x) = \boldsymbol{w} \cdot \boldsymbol{x_i} + b (2)$$

where $\boldsymbol{x_i} = (x_1,\ x_2,\ x_3, \dots, x_n)^T \in \mathbb{R}^n$, $\boldsymbol{w}$ denote the support vector weight value that perpendicular tits hyperplane, $\boldsymbol{x_i}$ the i-th data and b as the bias value. The solution of regression function in equation (2) can be done by minimizing the value of $\boldsymbol{w}$ as an optimization problem. Those can be solve by convex optimization problem by,

minimize

$$\frac{1}{2}\ \|\boldsymbol{w}\|^2\ (3)$$

Subject to

$$y_i - \boldsymbol{w} \cdot \boldsymbol{x_i} - b\ \leq \varepsilon$$
$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b - y_i\ \leq \varepsilon$$

In equation (3), there is a situation where the error value exceeds the threshold ε. The problem can be solved by using soft margins by adding slack variable to its margins $\xi_i\ , \xi_i^*$ to overcome the problem of infeasible constraints in optimisation problems by giving a penalty for data that does not fulfil the constraint. The penalty is notated with C, so the above formula can be changed into the following formula as follow.

minimize

$$\frac{1}{2}\ \|w\|^2 + C\ \sum_{i=0}^{n}(\xi_i + \xi_i^*) \qquad (4)$$

Subject to

$$y_i - \boldsymbol{w} \cdot \boldsymbol{x_i} - b\ \leq \varepsilon +\ \xi_i$$
$$\boldsymbol{w} \cdot \boldsymbol{x_i} + b - y_i\ \ \leq \varepsilon +\ \xi_i^*$$
$$\xi_i\ , \xi_i^*\ \geq 0 \quad \text{for } i\ = 1, 2, 3, \dots\ \ , i$$

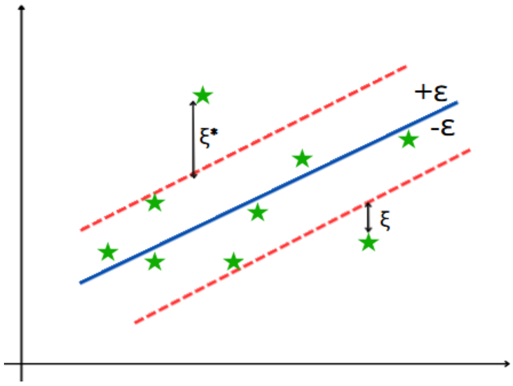

Figure 2. Soft Margin SVR Concept

Figure 2 illustrate how the slack variable were added in the soft margin concept, it is add because of there are some data that belong to other class but its in the hyperplane boundary. The solution for the optimization problem ini equation (4) is done by using the Langrange multiplier as follow.

$$L = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}(\eta_i\xi_i + \eta_i^*\xi_i^*) - \sum_{i=1}^{n}\alpha_i(\varepsilon + \xi_i - y_i + (\mathbf{w}.\mathbf{x_i})b) - \sum_{i=1}^{n}\alpha_i^*(\varepsilon + \xi_i^* + y_i - (\mathbf{w}.\mathbf{x_i}) - b) \quad (5)$$

The derivation formula process, the SVR regression function formula for the linear case is obtained as follows.

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)(x_i.x_j) + b \quad (6)$$

As for the nonlinear problem SVR function, it provides an alternative approach by mapping the data x from input space to feature space with higher dimension through a function φ such that $\varphi : x \rightarrow \varphi(x)$. Therefore, the nonlinear SVR regression function can be written as follows:

$$f(x) = \sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)\,K(x_i, x_j) + b \quad (7)$$

Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a global optimization method introduced by Kennedy and Eberhart in 1995 based on the behavior of flocks of birds and fish. Each particle in Particle Swarm Optimization has a particle movement speed in the search space with a speed that automatically adjusts to its previous behavior. Therefore, particles tend to move to better search regions during the search process [17]. PSO performs search using a population (swarm) of individuals (particles) that will be updated from one iteration to another. The PSO method can be assumed to be like a flock of birds looking for food in an area, but the flock does not know the exact location of the food. The following are the steps of the PSO algorithm process.

Initialization particle position

Assumed that a particle for each dimension must be in the domain defined by two vectors, x_{min} and, x_{max}, x_{min} represents the lower bound of each dimension and, x_{max} represents the upper bound of each dimension. The initialisation method to determine the particle position can be solved using the following equation:

$$x_{ij} = x_{min} + rand(x_{max} - x_{min})$$

x_{ij} = particle position; x_{min} represented the lower bound, x_{max} represented the upper bound and $rand$ represented the random value in range [0,1].

10.12928/bamme.v3i2.9561

Initialization particle speed

The initial velocity of the particle is initialized to 0 i.e., $v_i(0) = 0$. In this process it is necessary to limit the minimum and maximum velocities to prevent the particle from moving too far outside the search space. The initialization method for the speed limiter can be solved using the following equation.

$$v_{max,j} = k \frac{x_{max,j} - x_{min,j}}{2}$$

$v_{max,j}$ represented maximum particle velocity at dimension j, k represented the random value in the range [0,1], $x_{max,j}$ for upper limit of particle value in dimension j, and $x_{min,j}$ lower limit of particle value in dimension j.

Particle velocity has a speed limit or threshold that can be calculated using the following equation:

i. If $v_{ij}^{t+1} > v_{max,j}$, then $v_{ij}^{t+1} = v_{max,j}$
ii. If $v_{ij}^{t+1} < -v_{max,j}$, then $v_{ij}^{t+1} = -v_{max,j}$

Compute and evaluate the fitness

The fitness evaluation of the position can be calculated using the equation below:

$$fitness = \frac{1}{1 + MSE}$$

Determine the P_{best} and G_{best}

P_{best} is the best position ever reached by a particle. The process of finding the P_{best} value is done by comparing the P_{best} value with the current iteration particle. Meanwhile, G_{best} is the best position of the particle. The process of finding the G_{best} value is done by comparing the G_{best} and P_{best} values. If the G_{best} value is smaller than the P_{best} value, then the G_{best} position value remains. However, if the G_{best} value is greater than the P_{best} value, the G_{best} position value is replaced with the P_{best} location value.

Update the velocity and position values of the particle

The process of updating the velocity value and particle position is done using the following equation:

$$v_{ij}^{t+1} = \omega_t v_{ij}^t + c_1 r_{1j}^t \left(P_{best,ij}^t - x_{ij}^t\right) + c_2 r_{2j}^t (G_{best,ij}^t - x_{ij}^t)$$

$$x_{ij}^{t+1} = x_{ij}^t + v_{ij}^t$$

Checking the result

Checking the results whether the optimization solution from the process run is optimal or not. If the result obtained has reached the optimal result, the algorithm will stop. However, if the optimization solution has not been obtained or the maximum number of iterations is not reached, then repeat step 4 until the optimum solution is obtained or the maximum number of iterations is met.

Results and discussion

The data used in this study is the closing gold price data on 1 January 2021 to 31 January 2023. The data is obtained from the Yahoo!Finance website. Furthermore, data preprocessing is carried out by converting raw data into a form of data that is easier to process or understand. The purpose of this preprocessing process is to prepare the data before it is inputted into the model. This preprocessing consists of detecting missing data, data containing NaN, feature selection and data scaling. After preprocessing the data, we divide the data into training data and testing data. Training data is used to evaluate the strength or accuracy of the model. Testing data is data used to train the model and determine the relationship between features (independent variables) and targets (dependent variables). In this study, the dataset is divided into 70% training data and 30% testing data. The results of the dataset division obtained 364 training data and 157 testing data.

The next process is to form an SVR model, in forming an SVR model it is necessary to determine several parameters. These parameters include the C parameter, epsilon (ε) and gamma (γ). Determination of the parameters to be used in this study using the grid search algorithm process. The results of the parameter determination process using the grid search algorithm obtained $C = 10$, $\varepsilon = 0.01$ and $\gamma = 0.01$. The prediction process is carried out using these parameters and the RBF kernel using python software. Thus, the following prediction results are obtained.

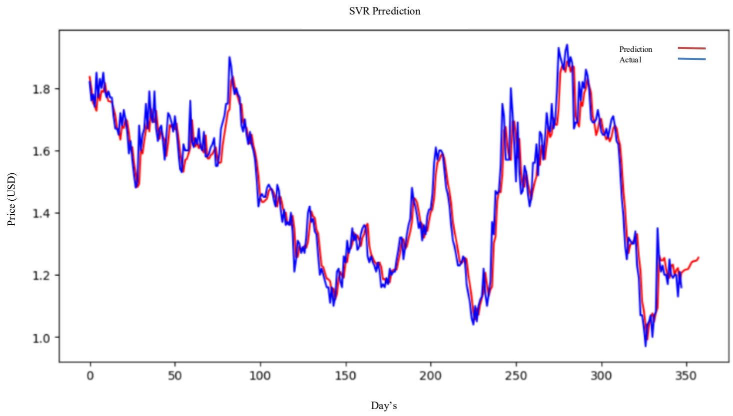

Figure 3. SVR Prediction Results Diagram

In Figure 3, the graph with the red color line is the pattern of the prediction results of the closing gold price. Meanwhile, the blue line is the actual gold cover price data pattern used to clarify the accuracy of the prediction through the intersection that occurs between the two data. Based on the graph, it shows that the SVR model with the tested parameters has a fairly good prediction ability. This is because the intersection of the graph between the actual data and the prediction almost coincides at several points.

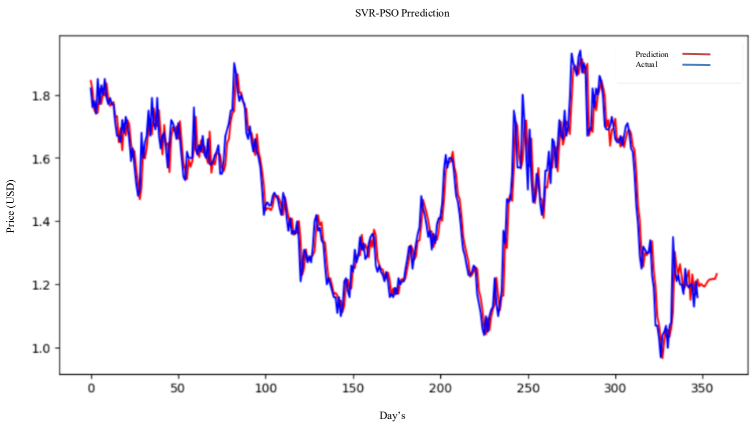

Figure 4. SVR-PSO Prediction Results Diagram

10.12928/bamme.v3i2.9561

Next, we will search for the optimal values of parameters C, epsilon (ε) and gamma (γ) for the SVR method using the PSO method. In this PSO process, the particle position and velocity initialisation process is carried out, calculating and evaluating the fitness value of each particle based on its position, each particle has a fitness value, determining P_{best} and G_{best}, and updating the velocity value and position value of the particle. So that the optimal parameters are C = 10000, ε = 0.035, and γ = 0.001. The prediction process is carried out using these parameters with the help of python software. Thus, the following prediction results were obtained.

In Figure 4, the graph with the red color line is the pattern of gold price prediction results. Meanwhile, the blue line is a pattern of actual closing gold price data used to clarify the accuracy of the prediction through the intersection that occurs between the two data. Based on the graph, it shows that the SVR-PSO model with the tested parameters has better prediction capabilities compared to predictions using the SVR model.

Next, the MSE calculation is carried out. The MSE calculation results obtained are presented in Table 1.

Table 1. Mean Square Eror Result for SVR and SVR - PSO

Method	MSE
SVR	0.0035744
SVR-PSO	0.0033058

Based on the results of the MSE and accuracy calculations in Table 1, it is obtained that the SVR and SVR with PSO methods have different MSE values. In the SVR method, the MSE value is 0.0035744. Meanwhile, the SVR method with PSO obtained an MSE value of 0.0033058. So, it can be concluded that there is a decrease in the MSE value of 0.0002686. Thus, it can be said that the SVR method with PSO is able to improve prediction performance compared to using the SVR method.

Conclusion

Based on the results and result decision, the SVR method with the PSO algorithm increase the accuracy and MSE values are smaller compared to the ordinary SVR. The MSE value obtained from each model is in the SVR model MSE value of 0.0035744. While the SVR with the PSO algorithm obtained an MSE value of 0.0033058. The use of the PSO (Particle Swarm Optimization) algorithm in the Support Vector Regression (SVR) model can optimize the SVR parameters, namely the C, epsilon (ε) and gamma (γ) parameters by finding the best combination or position to produce the minimum error. In addition, the PSO algorithm based on the concept of particle motion can explore the search space adaptively or the ability of particles to adjust their position and speed based on the $\boldsymbol{P_{best}}$ and G_{best} values so that optimal parameters can be achieved and produce better performance and prediction results.

Acknowledgement

Special thanks to our lecturer and friends for the discussion, especially.

References

Alaidi, A. H. M., Alairaji, R. M., Alrikabi, H. T. S., Aljazaery, I. A., & Abbood, S. H. (2022). Dark Web Illegal Activities Crawling and Classifying Using Data Mining Techniques. *Int. J. Interact. Mob.*

Technol., *16*(10), 122–139. https://dx.doi.org/10.3991/ijim.v16i10.30209

Ali, R., Mangla, I. U., Rehman, R. U., Xue, W., Naseem, M. A., & Ahmad, M. I. (2020). Exchange rate, gold price, and stock market nexus: A quantile regression approach. *Risks*, *8*(3), 1–16. https://dx.doi.org/10.3390/risks8030086

Ensafi, Y., Hassanzadeh, S., Zhang, G., & Shah, B. (2022). International Journal of Information Management Data Insights Time-series forecasting of seasonal items sales using machine learning – A comparative analysis. *Int. J. Inf. Manag. Data Insights*, *2*(1), 100058. https://dx.doi.org/10.1016/j.jjimei.2022.100058

Khaire, U. M., & Dhanalakshmi, R. (2022). .Stability of feature selection algorithm: A review. *J. King Saud Univ. Comput. Inf. Sci.*, *34*(4), 1060–1073. https://dx.doi.org/10.1016/j.jksuci.2019.06.012

Kim, M., Lee, J., Lee, C., & Jeong, J. (2022). Framework of 2D KDE and LSTM-Based Forecasting for Cost-Effective Inventory Management in Smart Manufacturing. *Appl. Sci.*, 12(5). https://dx.doi.org/10.3390/app12052380

Kusumawati, Y., Widyatmoko, K., & Irawan, C. (2022). Gold Price Prediction Using Support Vector Regression. *J. Appl. Intell. Syst.*, *7*(1), 89–102. https://dx.doi.org/10.33633/jais.v7i1.6124

Li, J. (2022). Macro carbon price prediction with support vector regression and Paris accord targets. September 2022, 1–13.

Madani, M. A., & Ftiti, Z. (2022). Is gold a hedge or safe haven against oil and currency market movements? A revisit using multifractal approach. *Ann. Oper. Res.*, *313*(1), 367–400. https://dx.doi.org/10.1007/s10479-021-04288-6

Maori, N. A. (2017). Perbandingan Metode ANN-PSO dan ANN-GA untuk Peningkatan Akurasi Prediksi Harga Emas ANTAM. *Disprotek*, *8*(1), 67–80.

Ngoc, T. T., Van Dai, L., & Thuyen, C. M. (2021). Support vector regression based on grid search method of hyperparameters for load forecasting. *Acta Polytech. Hungarica*, *18*(2), 143–158. https://dx.doi.org/10.12700/APH.18.2.2021.2.8

Rico, S., Ilham, N., Irada, S., & Mangasi. (2022). The Effect of Technical Analysis on Cryptocurrency Investment Returns With the 5 (Five) Highest Market. *Ejournal Sean Inst.*, *11*(2), 1022–1035.

Romana, A. S. (2017). A Comparative Study of Different Machine Learning Algorithms for Disease Prediction. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, *7*(7), 172. https://dx.doi.org/10.23956/ijarcsse/v7i7/0177

Setyowibowo, S., As'ad, M., Sujito, S., & Farida, E. (2022). Forecasting of Daily Gold Price using ARIMA-GARCH Hybrid Model. *J. Ekon. Pembang.*, *19*(2), 257–270. https://dx.doi.org/10.29259/jep.v19i2.13903

Syahri, A. & Robiyanto, R. (2020). The correlation of gold, exchange rate, and stock market on Covid-19 pandemic period. J. *Keuang. dan Perbank.*, *24*(3), 350–362. https://dx.doi.org/10.26905/jkdp.v24i3.4621

Tanveer, M., Rajani, T., Rastogi, R., Shao, Y. H., & Ganaie, M. A. (2022). Comprehensive review on twin support vector machines. *Ann. Oper. Res.* https://dx.doi.org/10.1007/s10479-022-04575-w

Vasista, K. (2022). Types and Risks Involved Towards Investing in Mutual Funds. *Int. J. Curr. Sci.*, *12*(1), 2250–1770.

Zhang, H. (2022). Unravelling the synergy of oxygen vacancies and gold nanostars in hematite for the electrochemical and photoelectrochemical oxygen evolution reaction. *Nano Energy*, *94*(January), 106968. https://dx.doi.org/10.1016/j.nanoen.2022.106968