Available online at : http://josi.ft.unand.ac.id/

Jurnal Optimasi Sistem Industri

| ISSN (Print) 2088-4842 | ISSN (Online) 2442-8795 |

Research Article

Optimizing Demand Forecasting Method with Support Vector Regression for Improved Inventory Planning

Tryantomo Lokhilmahful Palgunadi [a], Rina Fitriana [a,*], Anik Nur Habyba [a], Yun-Chia Liang [b]

[a] Department of Industrial Engineering, Universitas Trisakti, Jakarta, Indonesia
[b] Department of Industrial Engineering and Management, Yuan Ze University, Taoyuan, Taiwan

* Corresponding Author: rinaf@trisakti.ac.id

DOI: 10.25077/josi.v23.n2.p149-166.2024 **Submitted** : June 3, 2024; **Accepted** : November 16, 2024; **Published** : January 30, 2025

ABSTRACT

Problems arising from suboptimal production planning can cause inventory management to be less effective and efficient in the company. The lack of integrated presentation of information also causes less efficiency in making decisions. This study aims to obtain the best kernel function forecasting model by predicting ground rod sales using the Support Vector Regression (SVR) method in order to determine the level of forecasting accuracy and the results of ground rod forecasting in the future which are presented in an optimal data visualization. This problem-solving is done with the Support Vector Regression method, which consists of linear kernel functions, polynomial kernel functions, and radial basis function (RBF) kernel functions with the Grid Search Algorithm. Based on the results of the best parameter search that has been done using the grid search algorithm, it can be concluded that the best kernel function forecasting model is a linear kernel function with a value of C = 100 and $\varepsilon = 10^{-3}$. The accuracy of this forecasting model has a MAPE value of training data and testing data of 2.048% and 1.569%, where this value is the smallest MAPE value compared to the MAPE value of the other two functions. After getting the best model, forecasting was carried out within five months, obtaining an average of 6,647 monthly pieces. The results of forecasting and historical sales are reviewed in a visualization of Business Intelligence data so that it is well exposed, where the forecasting shows an increase from every month.

Keywords: Support Vector Regression (SVR), forecasting, grid search algorithm, Kernel functions, business intelligence

INTRODUCTION

In the development of the current technological age, the manufacturing industry has been considered one of the critical factors in developing countries, just like the manufacturing process both domestically and internationally in the industrial sector that has higher competition in meeting consumer demand with a reasonably volatile market. Therefore, production and inventory efficiency are critical aspects of the operations of a manufacturing company that affect overall performance. Demand forecasting plays an important role as a decision-support tool in industrial system management [1]. This study emphasize effective forecasting supports broader industrial objectives, like efficiency in resource use and operational sustainability [2]. Efficiency in inventory management is essential to control production costs and maintain adequate product availability for consumers. In the context of competitive business and fluctuating markets, manufacturing companies such as XYZ Co. are faced with demands to optimize their inventory management to achieve market production targets, especially in ground rod products. Demand for goods by customers tends to change and be uncertain. From 2021 to 2023, it can be seen that sales for ground rod

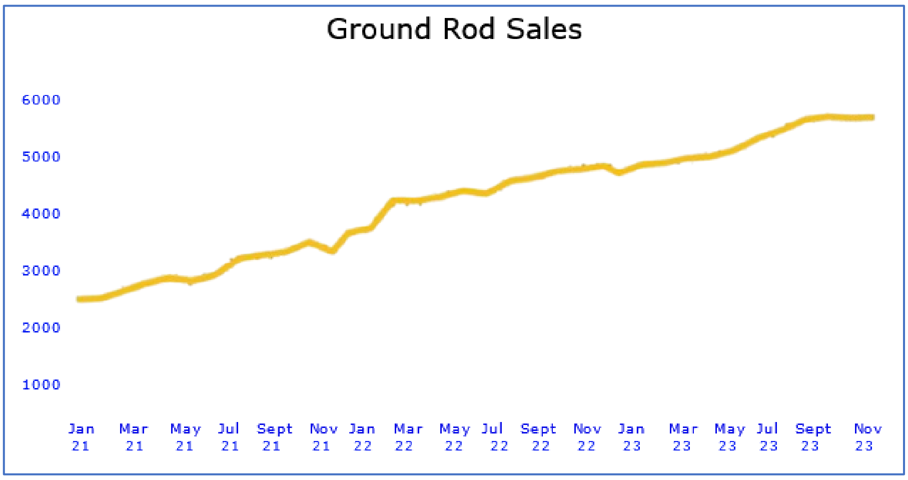

Figure 1. Ground rod sales graph in 2021-2023.

products fluctuates, but there tends to be an increase (Figure 1). The company XYZ Co. often finds it difficult to predict future demand, so this also results in less efficient ground rod stock. Starting from the ineffectiveness of forecasting methods, the problems resulting from this problem, apart from ineffective inventory stock, also impact production costs, overtime costs, and raw material ordering costs.

In addition to optimizing inventory management, manufacturing companies must also be able to produce products with qualities that follow national standards. This is done so that manufacturing companies are concerned not only with quantity or quality but because these two things are essential for the company's sustainability. Quality management is essential for companies because it directly impacts customer satisfaction and loyalty [3]. Thus, the company has noticed that quality is critical for developing products and services to drive sustainable achievement in previous research [4]. The sales they get results vary greatly, which means that different subjects are likely to have different data values for the same variable, and thus, the data values vary from subject to subject, which is the same as the previous study [5]. These changes in demand significantly affect the ground rod inventory owned by the company. Determination of ground rod inventory is done by comparing the number of items sold in the last month with the number of items that came out in the current month. This condition can impact production quality later, so if too much demand is not as predicted, the performance in producing ground rods will decrease [6]. The forecasting function is used to forecast the inventory of goods to be more precise based on past and future conditions. Forecasting is the science of predicting every event that will occur using past or historical data, which is processed and then projected into the future with a mathematical model [7]. So, based on these problems, this study uses forecasting methods to predict the inventory of goods that the company must prepare in the future. The method that can be used in this study is the quantitative forecasting method, which is a forecast based on quantitative data or mathematical models that vary with past data. In quantitative forecasting methods, there are three types of time-based forecasting methods, namely long-term (> two years), medium-term (3 months – 2 years), and short-term (0 – 3 months) [8]. In this case study, the forecasting method is the Support Vector Regression (SVR).

Machine learning methods such as Support Vector Regression (SVR) and kernel functions have been proposed to replace statistical methods in more recent sales predictions [9]. SVR has advantages in handling nonlinear problems with high dimensions and small samples [10]. It is based on statistical learning theory and structural risk minimization principles. SVR has become a popular and effective load prediction method in recent years. This is especially noticeable in research and practice, especially in operational research, where linear optimization techniques led in operational research are used to find optimal solutions to nonlinear prediction problems in larger-

 DOI:10.25077/josi.v23.n2.p149-166.2024

dimensional feature spaces [11]. The research used Support Vector Regression; we were able to model prediction models well for COVID-19 cases [12]. Then research [13][14] was also carried out using the SVR method to predict exchange rates/currency indices and stock indices so that SVR can be implemented in many objects, one of which is the sale of XYZ Co. products. This research approach was unique because combines the Support Vector Regression (SVR), the grid search algorithm, and power business intelligence for visualization data, bringing a new level of accuracy or efficiency to forecasting that wasn't addressed in previous studies.

This manufacturing company had problems planning Ground Rod production due to the lack of implementation of the forecasting method used. The company had only used a track record of every demand for ground rod products in the previous period, where the results still need to be optimal for predicting demand in the future. One of the impacts caused by this problem is the increase in company costs in providing overtime for employees. This research was conducted to optimize forecasting methods using support vector regression at XYZ Co. By proposing the forecasting model used, the company is expected to increase efficiency in ground rod product inventory management and minimize production costs at XYZ Co. Display visualization, recapitulation and data planning can also help companies carry out actions and decisions to maintain stability.

Forecasting is a process of forecasting future conditions using data from the past. Demand forecasting is essential for many companies because it determines production planning, inventory, and many other aspects of operations [9]. By using demand forecasting, the risks that occur in inventory control can be minimized. Demand forecasting helps businesses make decisions. The prediction of goal demand is divided into short-term and long-term. Short-run examples include production policies, pricing, sales controls, and finance, while long-term examples include determining production capacity [15]. A critical aspect of choosing a forecasting method is to look at patterns from the observed data. The prediction introduces four kinds of patterns: [16] cyclist pattern, horizontal pattern, trend pattern, and seasonal pattern. Support Vector Machine (SVM) is a mighty machine learning algorithm developed for classification and prediction problems, which works by recognizing patterns through kernel tricks. Due to its high performance and excellent generalizability compared to other classification methods, SVM methods are widely used in bioinformatics, text and image recognition, and finance, and one of them is prediction. This method finds the linear boundary (hyperplane) representing the most significant margin between two classes (labels) in the input space. This method can be applied not only to linear data but also to nonlinear data by using kernel functions [17]. Although first created to solve classification problems, regression work can also be handled using predictive analytics [18]. Support Vector Regression is an adaptation of support machine learning based on the classification of regression models from Support Vector Machine (SVM) [19]. SVR is a method that can solve nonlinear estimation problems. SVR successfully predicts time series data as well as time series financial data. According to Cortez and Vapnik [20], SVR uses linear models to be applied to nonlinear classes by mapping the input vector x into a high-dimensional feature space. SVR works by finding the maximum margin of the hyperplane and looking for the most minor error. The data closest to the maximum margin of the hyperplane is called the support vector [21]. Suppose the visualized data plot is feasible or feasible. In that case, a perfect regression will be obtained. However, suppose there is inappropriateness or $\varepsilon = 0$ infeasibility, where several points are out of the range or threshold. In that case, slack variables ξ and ξ^* will be added to optimize existing errors. This explanation can be seen in Figure 2.

In general, the SVM equation is shown by equation [21]:

$$f(x) = w^T \varphi(x) + b \tag{1}$$

where w is the weight factor, b is the threshold, and $\varphi(x)$ is a function that maps input x into the feature space. To minimize the defined risk function can be done by estimating the coefficients w and b with the following equation:

$$min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{\lambda} (\xi + \xi^*) \tag{2}$$

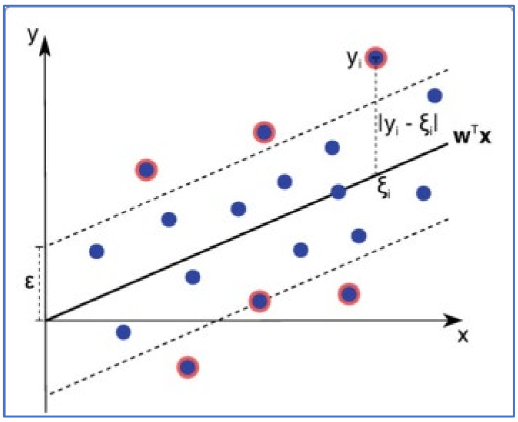

Figure 2. Support vector regression

Subject to:

$$y_i - \langle w\varphi(x_i) \rangle - b \leq \varepsilon + \xi \quad (3)$$

$$\langle w\varphi(x_i) \rangle + b - y_i \leq \varepsilon + \xi^* \quad (4)$$

$$\xi, \xi^* \geq 0 \quad (5)$$

In the equation above, the constant C has the conditions $C > 0$, which determines the trade-off between the flatness of the function f and the limit ε (epsilon). At the same time, the factor $\|w\|^2$ is also called regularization [20]. Any point or coordinate value outside the limit will be multiplied by C. Assigning a significant value of C means emphasizing the factor's importance over the function's flatness factor. Figure 3 graphically illustrates the situation by adding slack $\xi^*$and modifiers ξ. The slack variable located in the interval will be zero, and the farther the distance of the slack variable to the limit, the greater the value. Support vectors are points in the training data located exactly below and outside the boundary line. The smaller the accuracy value, the higher the number of ε support vectors, and the higher the value of the slack modifier.

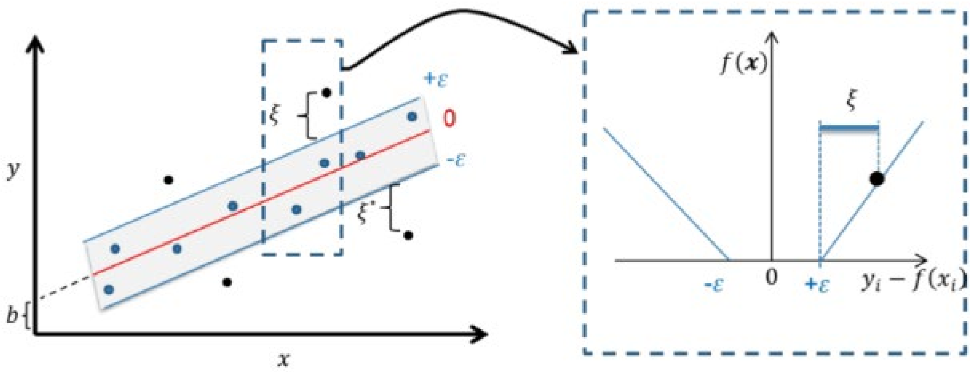

Figure 3. Random modifier illustration.

To solve this regression problem, the internal product of equation (1) can be replaced with the kernel function K(). This allows performing such operations in larger dimensions with low-dimensional space data input without knowing the transformation φ [23], as shown in equation (6). These are called kernel functions:

 DOI:10.25077/josi.v23.n2.p149-166.2024

$$f(x) = \sum_{i=1}^{N} (\alpha^* - \alpha). K(x_i, x_j) + b \tag{6}$$

The optimal squared problem associated with binder Lagrange α^* and α. SVR uses a kernel to connect data to a more prominent feature space, enabling more structured data processing and more efficient computation. Kernel functions in SVR have more precision than standard linear SVR in nonlinear regression [22]. This study uses three types of kernel functions: linear kernels, polynomial kernels, and Radial Basis Function (RBF) kernels. The formula of the three Kernel functions is as follows [23]:

- Kernel Linear
  $$K(x_i, x_j) = x_i^T x_j \tag{7}$$
  with the pairing of two data in the x_i, x_j training data
- Kernel Polynomial
  $$K(x_i, x_j) = (x_i^T x_j + 1)^d, d > 1 \tag{8}$$
  where d is the degree of the required parameter
- Kernel Radial Basis Function
  $$K(x_i, x_j) = exp\left(-\gamma \left\| x_i - x_j \right\|\right)^2 \tag{9}$$

METHODS

Figure 4 explains the stages in this research, which include problem identification, literature studies and field studies, data collection, data preprocessing, parameter tuning with grid search, SVR forecasting with kernel functions, validation with MAPE and R^2, predicting ground rod requests, and also visualizing data with Power BI.

Problem Identification

At this stage, an identification of the problems to be solved in this research will be identified. This study's problem is the demand for ground rod products at XYZ Co. This research will forecast sales or demand for ground rods at XYZ Co. to prevent overtime, additional production costs, or other costs.

Literature Studies and Field Studies

At this stage, several literature reviews will be discussed that support the work of this research. The literature used includes international journals, books, and national journals that dissect the topics determined in this study. At this stage, the results are research concepts and methods, namely the SVR method with the Grid Search Algorithm to find optimal parameters.

Data Collection

At this stage is to collect data. The data collected is ground rod sales data from XYZ Co. from 2021 to 2023.

Preprocessing Data

This stage is the stage after the data is obtained. This stage is the data preprocessing stage, where the data used for making models, namely XYZ Co. ground rod sales data, will be divided into training data and testing data with a proportion of 80% of the total observation data for training data and 20% observation data for test data. From the total data, 36 data, 28 training data, and eight testing data were obtained. Before the separation between test data and training data, preprocessing was carried out in the form of analyzing raw data by mapping and grouping data into periods per month, then checking null descriptive data from all ground rod product observation data at XYZ Co., which was processed using Python language.

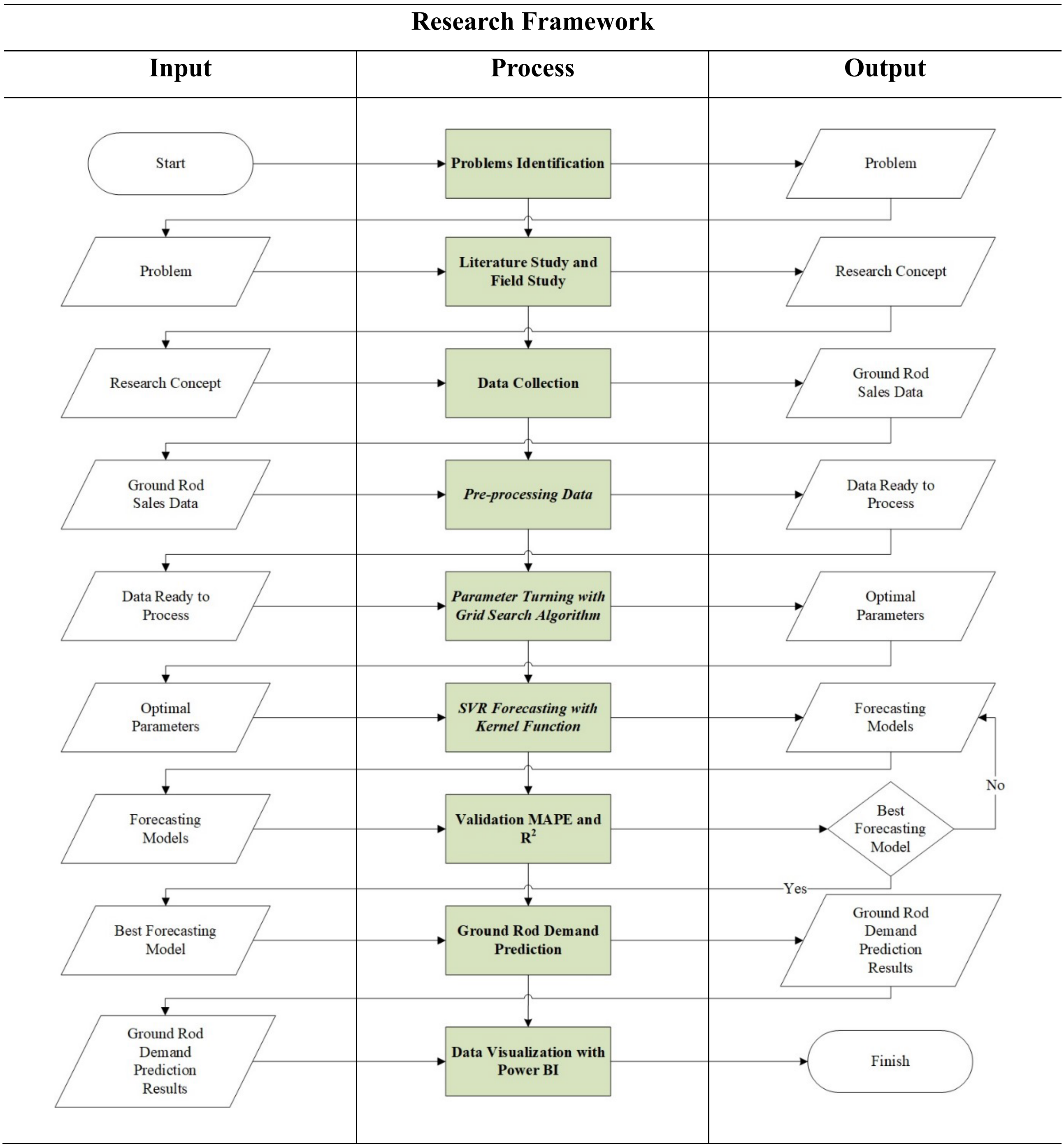

Figure 4. Research framework

Parameter Tuning with Grid Search Algorithm

The Grid Search algorithm can be used to find the ideal parameters. This Algorithm divides the range of parameters to be optimized into grids and traverses each point to obtain the ideal parameters. In its applications, various performance metrics must control grid search algorithms, which are usually measured through cross-validation tests on training data. Therefore, it is recommended to try several parameter pairs on the SVR hyperplane [26]. The parameter pair that results in the most accurate cross-validation test is the best. The cross-validation method is used when selecting model combinations and hyperparameters that automatically validate each combination [26]. The optimal pair of hyperparameters has the most minor error value and the highest accuracy.

 DOI:10.25077/josi.v23.n2.p149-166.2024

Table 1. MAPE Descriptions

MAPE	Description
< 10%	The ability of the forecasting model is excellent
10% - 20%	The ability of the forecasting model is good
20% - 50%	The ability of the forecasting model is quite decent
> 50%	The ability of the forecasting model is terrible

Forecasting and Support Vector Regression

Forecasting uses the Support Vector Regression (SVR) algorithm with kernel functions. SVR is a machine learning algorithm used for forecasting and classification. Kernel functions convert input data into higher feature spaces to be separated appropriately. Some commonly used types of kernel functions are linear, polynomial, and radial basis functions (RBF). This study will determine the kernel functions used in SVR using these three kernel functions in Equation (7)-(9).

Validation Measurement with MAPE and R^2

The error is usually calculated using Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). The MAPE description is presented in Table 1. Mean absolute error is the average of the absolute values of forecasting errors. Mean squared error is the average of forecasting errors squared. The mean absolute percentage error is the percentage of error of experience [27]. However, MAPE will be used in this study. Here is the equation:

$$MAPE = \left[\frac{1}{n}\sum_{i=1}^{n}\left|\frac{e_i}{D_i}\right|\right] \times 100\% \tag{10}$$

where:

$e_i = F_t - D_t$	Forecast errors
F_t	Forecast value
D_t	Actual value
t	Period

The coefficient of determination (R^2) is an accuracy metric that indicates the magnitude of the dependent variable in the model [26]. It ranges from zero to one or zero to one hundred percent. A good coefficient of determination is a value close to one. Here is the equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(F_t - \overline{D_t})}{\sum_{i=1}^{n}(D_t - \overline{D_t})} \tag{11}$$

In the study [29], it was further explained as follows:

- $R^2 \geq 0$ means that with linear regression without constraints, R^2 is non-negative and corresponds to the square of the double correlation coefficient.
- $R^2 = 0$ means that the attached line (hyperplane) is horizontal. With two numerical variables, this happens if the variables are independent of each other (independent) or uncorrelated.
- $R^2 < 0$ means that this case is only possible with linear regression when the intercept or slope is delimited so that the "most suitable" line (given its constraints) is worse than the horizontal line, for example, if the regression line (hyperplane) does not follow the data. Table 2 shows the categories of R^2.

Table 2. R^2 Categories

R^2	Description
0,75 - 1	Strong Category
0,50 - 0,75	Moderate Category
0,25 - 0,50	Weak Category
< 0,25	Very Weak Category

Data visualization with Power Business Intelligence (BI)

Business Intelligence (BI) is the application of available mathematical models and data analysis methodologies to obtain relevant information and knowledge in the face of complex decisions. This system acts as a tool that provides accurate and valuable information to decision-makers within a specific time limit, thus supporting the decision-making process in the dairy industry [30]. Power BI is a cutting-edge business intelligence and data visualization platform developed by Microsoft. It empowers organizations and individuals to turn raw data into valuable insights and visually appealing reports and dashboards [31]. Power BI has advantages in ease of use, integration with various data sources, and flexibility in creating attractive visualizations. In addition, business intelligence can recognize unstructured data, improve product quality, and create new opportunities in business strategy [32]. The use of BI offers several benefits that can speed up decision-making, such as [33] creating dashboards and performance measurement metrics in the era of automation and acceleration. A focus on business attention on key customers and channels that drive desired results, measure performance against established performance measurement goals and can influence management with exclusion strategies.

RESULT AND DISCUSSION

Step 1: Preprocessing Data

After the data is obtained on the observation results, data processing is carried out. The data that serves to create the model is historical data on sales of XYZ Co. ground rod products, divided into training sets and test sets. The training data and test data are divided by the proportion of 80% as training data from all observation data and 20% as test data. For each variable, 28 training data and 8 test data were obtained from a total of 36 data obtained. Training data has more testing data so that machine learning models can learn, shape, and create the best models [34][35]. Before the separation between test data and training data, preprocessing was carried out in the form of analyzing raw data by mapping and grouping data into periods per month, then checking null descriptive data from all ground rod product observation data at XYZ Co. which was processed using python language which results are presented in Table 3.

Descriptive statistics provide an overview of the sample under study without concluding based on probability theory. Descriptive statistics provide a general overview, although the study's primary purpose is inferential statistics. Descriptive statistics describes a population using tools such as frequency tables, percentages, and other measures of central tendency such as mean [36]. The information provided in Table 3 is a descriptive statistical table containing several central measurements such as the amount of data (n), mean, minimum value, maximum value, and standard deviation. The mean is the arithmetic mean of the sum of the data values divided by the number of observations,

Table 3. Descriptive Statistics

n	Mean	Minimum	Maxsimum	Std. Deviation
36	4117.44	2611	5444	898.52

 DOI:10.25077/josi.v23.n2.p149-166.2024

one measure of central tendency. The standard deviation of a distribution can be used to describe its spread. In other words, standard deviation expresses how close each observed value is to the mean [37]. The information shows an interpretation that the most significant ground rod sales over the last three years experienced by XYZ Co. were 5444 pieces in December 2023, while the most miniature ground rods ever experienced by XYZ Co. were 2611 pieces in January 2021. Then, the average sales of ground rod products from XYZ Co. from 2021 to 2023 were obtained at 4117.4 or around 4117 pieces with a standard deviation of 898.52. Although demand or sales over the past three years at XYZ Co. are not easy to predict, the company has shown growth in sales of ground rod products over the last three years.

Step 2: Parameter Tuning with Grid Search Algorithm

Before building a prediction model with the Support Vector Regression (SVR) method, parameter tuning is done to get optimal parameters for building the model. The Grid Search algorithm is used to find the best combination of parameters by iterating over all possible combinations of predefined parameters [38]. Parameter tuning is carried out using the core parameters of each SVR algorithm. The process of obtaining optimal parameters of kernel functions is carried out using the grid search algorithm. In searching for optimal parameters, tuning parameters are used, which, as SVR input, are the values of C (cost), γ (gamma), d (degree), and ε (epsilon). SVR parameter search with a Grid search Algorithm divides the range of parameters to be optimized into grids and across each point to get the ideal parameter. Parameter tuning is done for each kernel using the Grid search CV function in the sklearn model selection library in Python software. The grid search must also be readjusted to the kernel that is being used. In linear kernels, grid search is used in finding optimal parameters of the constants C and ε (epsilon). While in the polynomial kernel, grid search is used to find and optimize parameters with constants C, ε, and d (degree). Then in the Radial Basis Function (RBF) kernel, grid search functions as searching and optimizing parameters with constants C, ε, γ (gamma). In determining the magnitude of the γ parameters influenced from each data point in the input space mapping to higher dimensions.

The values of the parameters used in the three kernel functions vary. Table 4 describes how the grid search algorithm is performed in stages. Starting from input in the form of time series data, which is then processed using a grid search by entering predetermined parameters, then the best parameter search process with 'gridsearch best estimator'. After obtaining the optimal parameters, predictions are made with the model (Mi) set to see the value of the evaluation metric (Mean Absolute Percentage Error). After MAPE validation has been performed and valid, the model (Mi) is ready to be used as a prediction model. Table 5 shows that in linear kernel functions, the C and ε

Table 4. Computational Design Grid Search Algorithm

Algorithm	: Algorithm design for grid search to find the best testing models.
Input	: training time series (*R*); testing time series (*T*)
Output	: the best parameters with minimum testing error (*BTM*)

FindBestTestingModels (*R*, *T*)

- **foreach** C $\in \{10^{-1}, 1, \ldots, 100\}$; ε $\in \{10^{-3}, 10^{-2}, 10^{-1}\}$; $\gamma \in \{10^{-5}, 10^{-4}, \ldots, 10^{-1}\}$; d $\in \{2, 4\}$ **do**
  - P ← Use the best parameter grid estimator.
  - M_i ← Make the i-th model with parameters P
  - Train M_i with R
  - F_i ← Predict using the trained model M_i
  - $MAPE_i$ ← MAPE between the predicted (F_i) and the original testing data (F_i)
  - BTM_L ← add the best model M_i
- **return** BTM

Table 5. Kernel Function Parameter Values

No.	Kernel Functions	Parameter Hyperplane	Parameter Value
1	Kernel Linear	C	10^{-1}, 1, 10, 100
		ε	10^{-3}, 10^{-2}, 10^{-1}
2	Kernel Polynomial	C	10^{-1}, 1, 10, 100
		ε	10^{-3}, 10^{-2}, 10^{-1}
		d	2, 4
3	Kernel Radial Basis Function (RBF)	C	10^{-1}, 1, 10, 100
		ε	10^{-3}, 10^{-2}, 10^{-1}
		γ	10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}

parameters used are as high as 10-1, 1, 10, 100 and 10-3, 10-2, 10-1. Then, in the linear kernel function, the parameters C, ε, γ, and d used are C = 10-1, 1, 10, 100, then ε = 10-3, 10-2, 10-1, and d = 2, 4. Next, in the Radial Basis Function (RBF) kernel, the parameters C, ε, and γ used are C = 10-1, 1, 10, 100, ε = 10-3, 10-2, 10-1, and γ = 10-5, 10-4, 10-3, 10-2, 10-1. Of the various parameters that have been determined, the best parameter search process will be carried out using the grid search algorithm so that the model can run optimally on each kernel function.

Step 3: MAPE and R^2 Kernel Function Validation

After tuning the parameters using the grid search algorithm, the next step is to perform and run the model on each kernel to see which of each prediction will later be selected as the best prediction model. These parameters are passed into the SVR model in Python. After obtaining the prediction model, the results are interpreted in Table 6.

After searching for the best parameters using the grid search algorithm, it can be seen in Table 4 that the best parameter results obtained in linear kernel functions are C = 100 and ε = 10-3 with MAPE values of training data and testing data produced are 2.048% and 1.569% [37]. Then, the best parameter results obtained in polynomial kernel functions are C = 1, ε = 10-3, and d = 2. Training data and testing data produced are 2.705% and 2.430%. Furthermore, the Radial Basis Function (RBF) kernel function obtained the best parameter results with values of C = 100, ε = 10-3, and γ = 10-5 training data and testing data produced were 5.777% and 24.917%. Based on the results of the best parameter search that has been done using the grid search algorithm, it can be concluded that the best kernel function that can be used as a prediction model is a linear kernel function with a value of C = 100 and ε = 10-3, which has a MAPE value of training data and testing data by 2.048% and 1.569% where these values are the smallest MAPE values compared to MAPE values in polynomial kernel functions and radial base function kernel functions. Therefore, it can be said that linear kernel functions have prediction values with the smallest error rate compared to

Table 6. Best Parameter Values of Kernel Functions

No.	Kernel Function	Parameter Hyperplane	The Best Parameter Value	MAPE Value (%)	
				Training Data	Testing Data
1	Kernel Linear	C $(cost)$	100	2,048	1,569
		ε $(epsilon)$	10^{-3}		
2	Kernel Polynomial	C $(cost)$	1	2,705	2,430
		ε $(epsilon)$	10^{-3}		
		d $(degree)$	2		
3	Kernel Radial Basis Function (RBF)	C $(cost)$	100	5,777	24,917
		ε $(epsilon)$	10^{-3}		
		γ $(gamma)$	10^{-5}		

DOI:10.25077/josi.v23.n2.p149-166.2024

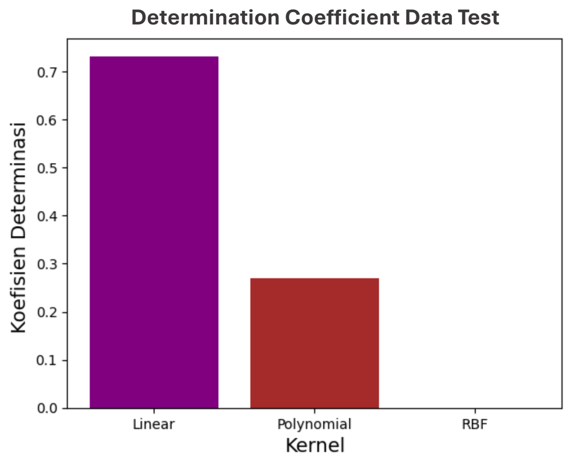

Figure 4. Comparison graph of Kernel data testing function

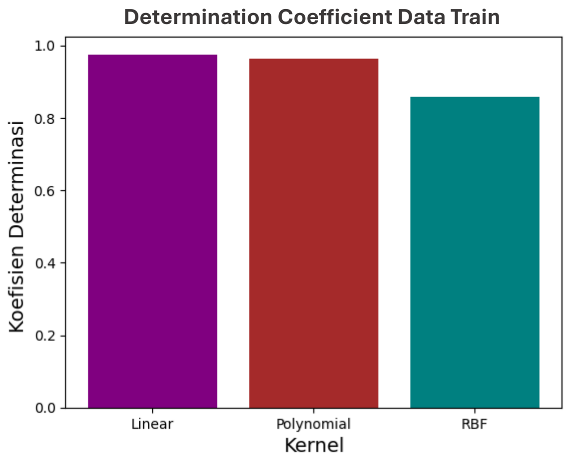

Figure 5. Comparison graph of Kernel data training function

polynomial kernel functions and RBF kernel functions. Verification using the best prediction model by looking at the three kernel functions' coefficients of determination (R^2). The use or function of the coefficient of the determination itself is to show the match of the estimated results of a regression model with the data modeled. In other words, the coefficient of determination (R^2) ensures the accuracy of the prediction model in regression analysis. In the R^2 calculation process, Python software is needed to calculate metrics from each Kernel function.

Based on Figure 4 and Figure 5, the coefficient of determination (R^2) value in each kernel function consists of two types, namely the value of the coefficient of determination of training data and the value of the coefficient of determination of testing data. From the two figures, it can be seen that in the training data, the most significant coefficient value falls on the linear kernel function, which is 0.975, while the value of the determination coefficient on the polynomial kernel function and the linear, radial base function kernel function is 0.964 and 0.858. Then, in the testing data, the most significant value of the coefficient of determination is also occupied by the linear kernel function, which is 0.732. At the same time, in the polynomial kernel function and the radial kernel function, the base function has values of 0.270 and -51.1. The value of the coefficient of determination in the training and test data shows that the linear kernel function has a better regression model accuracy for prediction than the other two functions. This is shown in the kernel radial base function, which has a negative coefficient of determination, which means that there is no influence of the model provided on the data presented [27].

Step 4: Ground Rod Sales Prediction

Using vector regression support and grid search algorithms, the best regression model is obtained, a linear kernel function model with C = 100 and $\varepsilon = 10^{-3}$. This model will then predict ground rod sales in the coming period. Before making predictions, this best model will show the results of the prediction graph both in the training and testing data, which will be shown in the visualization results in Figure 6 and Figure 7. The actual data and predictions that have been plotted show the slightest shift in predictions made by the model. This shows that the model can predict the ground rod sales level that will be used in the next period.

Table 7 shows the forecasting results of the selected regression model. Based on the table, sales or demand for ground rod products tend to increase, so XYZ Co. must prepare good production planning in order to meet customer demand. The recommendations for practitioners include using the linear kernel in similar demand forecasting applications to enhance accuracy and reduce costs.

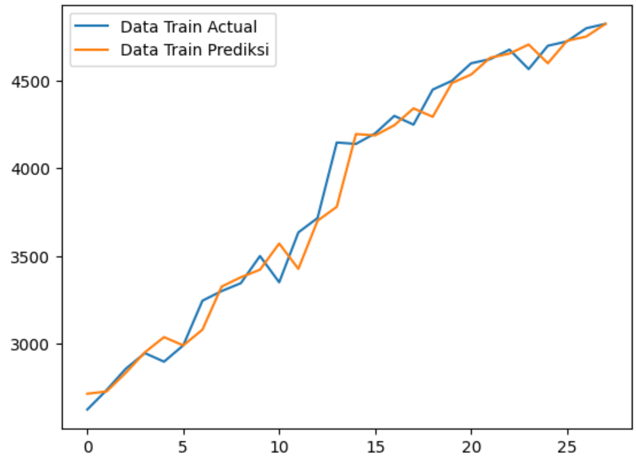

Figure 6. Graph of prediction data and actual data on training data training

Figure 7. Graph of prediction data and actual data on training data testing

Table 7. Ground Rod Sales Prediction Results for the Next 5 Months

Month	Forecast
1	5697
2	6058
3	6593
4	7444
5	7444
Total	33.236

Step 5: Data visualization with Power BI

After getting the best model results, the overall data results at XYZ Co. can be displayed interactively and interestingly by visualizing Power BI. The goal is that in addition to summarizing the overall results, it can also display them in an integrated, clear, and exciting way to present. The data visualization this time will be displayed consisting of the following:

 DOI:10.25077/josi.v23.n2.p149-166.2024

Sales by Product

Figure 8 explains that the ground rod products sold have several types, including ground rod products with a size of 1.5 meters (Ground Rod PK-AS 5/8 x 1500 mm), ground rods with a size of 2.4 meters (Ground Rod PK-AS 5/8 x 2400 mm), and ground rods with a size of 2.5 meters (Ground Rod PK-AS 5/8 x 2500 mm). The sales of these three types of ground rod products differ in the overall number sold in 2021 to 2023, for sales of ground rod products with a size of 1.5 meters (Ground Rod PK-AS 5/8 x 1500 mm) marked in purple has a sales contribution of 33,720 pieces or around 22.38% of the total sales. As for the sale of ground rods with a size of 2.4 meters (Ground Rod PK-AS 5/8 x 2400 mm) marked in blue has a sales contribution of 80,046 pieces, or around 52.73% of the total sales. Then, the sale of ground rods with a size of 2.5 meters (Ground Rod PK-AS 5/8 x 2500 mm) marked in pink has a sales contribution of 37,371 pieces or around 24.88% of the total sales. Thus, for the last three years, ground rod products with the most significant sales contribution are ground rod products with a size of 2.4 meters (Ground Rod PK-AS 5/8 x 2400 mm). In this dashboard, a filter feature is used both in years and months because this is useful so that companies can display directly what types of products can be seen within a certain period to conduct market analyses to increase profits and evaluation analysis for quality improvements for products that still have a small contribution.

Historical Graph of Ground Rod Product Sales

Figure 9 shows a graph depicting the overall sales of ground rod products from 2021 to 2023. For this dashboard, several features are used to see and monitor both for a predetermined month scale or year, which helps get information as needed in the timeline. This dashboard also shows the total results of all ground rod product sales in the last three years, which amounted to 148,228 pieces.

Ground Rod Product Sales/Demand Chart

Figure 10 shows a dashboard showing forecasting results using the linear kernel function's Support Vector Regression (SVR) method. As a result, in the next five months, demand for ground rod products tends to increase from January to May. For the total predicted demand for ground rod products in the next five months, 33,236 pieces were obtained.

Ground Rod Product Sales Dashboard Monitoring

Figure 11 describes the whole in terms of historical sales and forecasting later sales. In this dashboard, several features have been summarized from previous dashboards. These features consist of a time filter that can display sales from

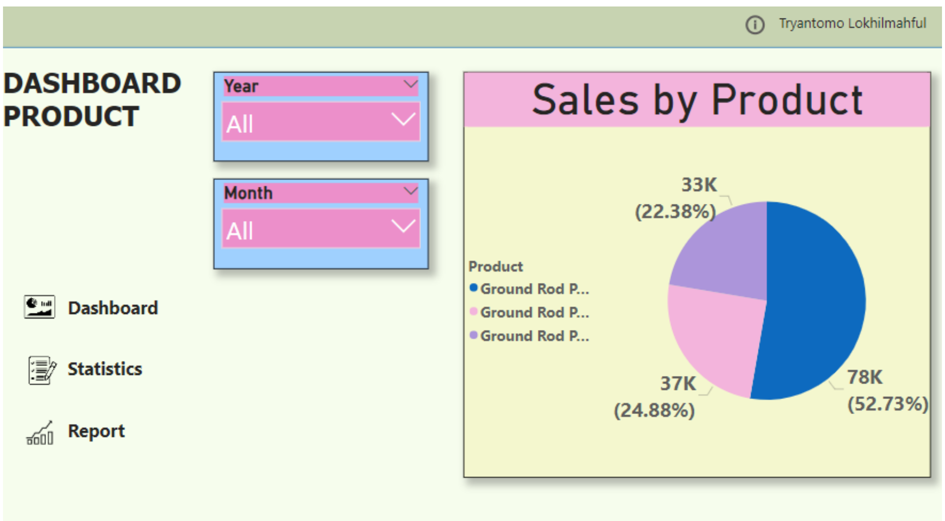

Figure 8. Dashboard product

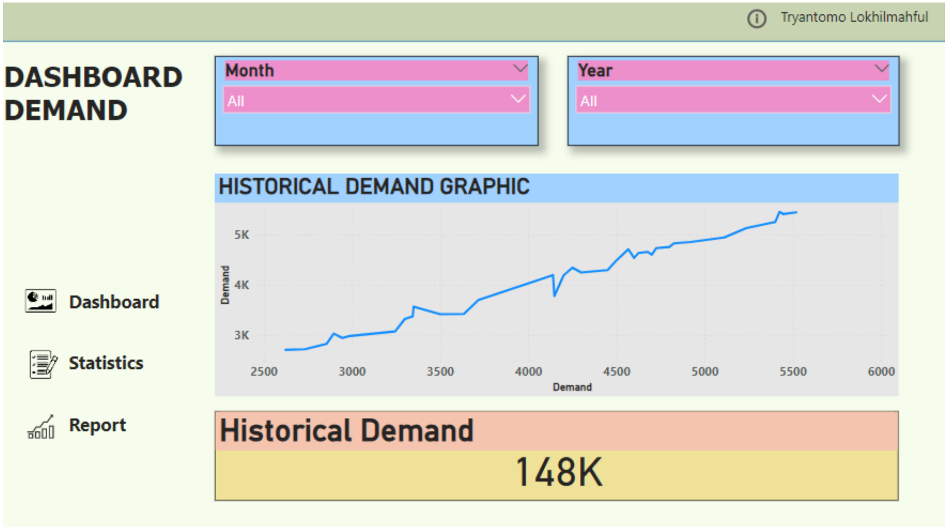

Figure 9. Dashboard demand

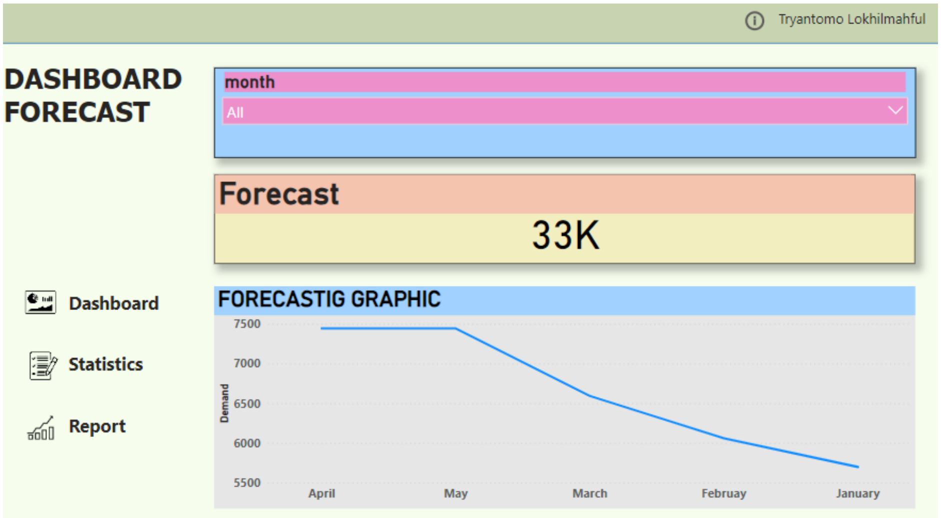

Fig 10. Forecast Dashboard

the cumulative side and the product type sold will appear according to a predetermined time. Then there is a pie chart that shows sales based on the type of product that has been sold, which can then be seen as the cumulative total by looking at the display of the "historical demand" feature according to the type of product chosen later. Furthermore, the last is an illustration in the form of a graph showing the track record of sales of ground rod products in the last three years, as well as forecasting of ground rod products in the next five months along with the cumulative total of the forecast.

CONCLUSION

The prediction model of the linear kernel and grid search algorithms is the most accurate model for XYZ Co.'s demand forecasting. The model optimizes parameters with MAPE values of training data and testing data. The proposed data visualization uses Power Business Intelligence software to display any data information. The data visualization proposal presented is a dashboard that produces several information presentations, including what has been created. There is a product dashboard that functions to recapitulate ground rod sales displayed based on the

 DOI: 10.25077/josi.v23.n2.p149-166.2024

type of product itself to see the highest contribution in getting the most significant profit. These results can benefit company production planning and inventory control processes by reducing costs, optimizing resource allocation, and enhancing overall efficiency.

ACKNOWLEDGMENT

We would like to express my sincere gratitude to the editor and the anonymous reviewers for their invaluable time, expertise, and constructive feedback. Their thoughtful comments and insightful suggestions significantly contributed to the improvement and refinement of this manuscript. I appreciate their commitment to maintaining the quality and rigor of the academic publication process.

CONFLICT OF INTEREST

The authors declare no conflicts of interest that could potentially compromise the objectivity, impartiality, or credibility of the research presented in this article.

FUNDING

The authors received no financial support for the research, authorship, and/or publication of this article.

References

[1] S. Fattaheian-Dehkordi, A. Fereidunian, H. Gholami-Dehkordi, and H. Lesani, "Hour-ahead demand forecasting in smart grid using support vector regression (SVR)," *International Transactions on Electrical Energy Systems*, vol. 24, no. 12, pp. 1650–1663, Sep. 2013. doi: 10.1002/etep.1743.

[2] E. Fradinata, Z. M. Kesuma, S. Rusdiana, and N. Zaman, "Forecast analysis of instant noodle demand using support vector regression (SVR)," *IOP Conference Series: Materials Science and Engineering*, vol. 506, p. 012021, Apr. 2019. doi: 10.1088/1757-899X/506/1/012021.

[3] R. A. Ramadhani, R. Fitriana, A. N. Habyba, and Y. C. Liang, "Enhancing quality control of packaging product: A Six Sigma and data mining approach," *J. Optimasi Sist. Ind.*, vol. 22, no. 2, pp. 197–214, 2023. doi: 10.25077/josi.v22.n2.p197-214.2023.

[4] M. Y. Latansa, T. J. Wahyu Adi, and M. A. Rohman, "Analysis the effect of total quality management, technology, and knowledge management on product quality in Indonesia construction industry," *IPTEK Journal of Proceedings Series*, vol. 0, no. 5, p. 527, 2019. doi: 10.12962/j23546026.y2019i5.8246.

[5] C. Shoko and C. Sigauke, "Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data," *Am. J. Infect. Control*, vol. 51, no. 10, pp. 1095–1107, 2023. doi: 10.1016/j.ajic.2023.05.009.

[6] A. D. De Menezes, J. G. M. De Barros, B. B. Da Fonseca, M. V. C. Aguilera, and N. A. de S. Sampaio, "Demand forecasting and inventory management as requirements for quality assurance in a bakery company," *Rev. Gestão e Secr. (Management Adm. Prof. Rev.)*, vol. 14, no. 9, pp. 15680–15694, 2023. doi: 10.18577/gestao.v14n9.2023.1247.

[7] O. Duru and M. Butler, "Modeling and forecasting with fuzzy time series and artificial neural networks," *Adv. Bus. Manag. Forecast.*, vol. 12, pp. 155–180, 2017. doi: 10.1108/S1477-407020170000012006.

[8] S. K. Selvam, C. Rajendran, and G. Sankaralingam, "A linear programming-based bi-objective optimization for forecasting short univariate time series," *Decis. Anal. J.*, vol. 10, p. 100400, 2024. doi: 10.1016/j.dajour.2023.100400.

[9] V. M. Evangelista and R. G. Regis, “Exploring the suitability of support vector regression and radial basis function approximation to forecast sales of Fortune 500 companies,” *Adv. Bus. Manag. Forecast.*, vol. 13, pp. 3–23, 2019. doi: 10.1108/S1477-407020190000013006.

[10] D. Indra Purnama and S. Setianingsih, “Model support vector regression (SVR) untuk peramalan jumlah penumpang penerbangan domestik di Bandara Sultan Hasanudin Makassar,” *J. Mat. Stat. dan Komputasi*, vol. 16, no. 3, pp. 391–403, 2020. doi: 10.20956/jmsk.v.

[11] S. Maldonado, A. González, and S. Crone, “Automatic time series analysis for electric load forecasting via support vector regression,” *Appl. Soft Comput. J.*, vol. 83, p. 105616, 2019. doi: 10.1016/j.asoc.2019.105616.

[12] D. Parbat and M. Chakraborty, “A python-based support vector regression model for prediction of COVID-19 cases in India,” *Chaos, Solitons and Fractals*, vol. 138, p. 109942, 2020. doi: 10.1016/j.chaos.2020.109942.

[13] H. Niu, “A method to increase the analysis accuracy of stock market valuation: A case study of the Nasdaq Index,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 1, pp. 732–743, 2024. doi: 10.14569/IJACSA.2024.0150173.

[14] T. Papadimitriou, P. Gogas, and V. Plakandaras, “Testing exchange rate models in a small open economy: An SVR approach,” *Bull. Appl. Econ.*, vol. 3, no. 2, pp. 9–29, 2016.

[15] A. Irankhah, M. H. Yaghmaee, and S. Ershadi-Nasab, “Optimized short-term load forecasting in residential buildings based on deep learning methods for different time horizons,” *J. Build. Eng.*, vol. 84, p. 108505, 2024. doi: 10.1016/j.jobe.2024.108505.

[16] B. H. Prakoso, “Implementasi support vector regression pada prediksi inflasi indeks harga konsumen,” *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 19, no. 1, pp. 155–162, 2019. doi: 10.30812/matrik.v19i1.511.

[17] M. Ryu and K. Lee, “Selection of support vector candidates using relative support distance for sustainability in large-scale support vector machines,” *Appl. Sci.*, vol. 10, no. 19, pp. 1–14, 2020. doi: 10.3390/app10196979.

[18] R. Zhang, Y. Li, Y. Gui, D. J. Armaghani, and M. Yari, “A stacked multiple kernel support vector machine for blast induced flyrock prediction,” *Geohazard Mech.*, Jan. 2024. doi: 10.1016/j.ghm.2024.01.002.

[19] B. M. Henrique, V. A. Sobreiro, and H. Kimura, “Stock price prediction using support vector regression on daily and up to the minute prices,” *J. Financ. Data Sci.*, vol. 4, no. 3, pp. 183–201, 2018. doi: 10.1016/j.jfds.2018.04.003

[20] E. S. Hedianti, “Peramalan harga saham dengan menggunakan metode support vector regression (SVR),” unpublished, 2019.

[21] D. Sepri and A. Fauzi, “Prediksi harga cabai merah menggunakan support vector regression,” *CBIS Journal*, vol. 8, no. 2, pp. 1–5, 2020.

[22] M. Awad and R. Khanna, *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Springer, 2015. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4302-5990-9.

[23] G. Astudillo, R. Carrasco, C. Fernández-Campusano, and M. Chacón, “Copper price prediction using support vector regression technique,” *Appl. Sci.*, vol. 10, no. 19, pp. 1–13, 2020. doi: 10.3390/APP10196648.

[24] R. Babagoli and M. Rezaei, “Prediction of moisture resistance of asphalt mastics modified by liquid anti-stripping based on support vector regression, artificial neural network and kernel-based support vector regression methods,” *Constr. Build. Mater.*, vol. 335, no. March, p. 127480, 2022. doi: 10.1016/j.conbuildmat.2022.127480.

[25] A. Hermawan, I. W. Mangku, N. K. K. Ardana, and H. Sumarno, “Analisis support vector regression dengan algoritma grid search untuk memprediksi harga saham,” *J. Math. Its Appl.*, vol. 18, no. 1, pp. 41–60, 2022. doi: 10.29244/milang.18.1.41-60.

[26] S. Xie, H. Lin, Y. Chen, H. Duan, H. Liu, and B. Liu, "Prediction of shear strength of rock fractures using support vector regression and grid search optimization," *Mater. Today Commun.*, vol. 36, no. May, p. 106780, 2023. doi: 10.1016/j.mtcomm.2023.106780.

[27] W. Yi, I. V. Gerasimov, S. A. Kuzmin, and H. He, "An intelligent algorithm of support vector regression parameters optimization in soft measurements," *Proc. 19th Int. Conf. Soft Comput. Meas. SCM 2016*, no. 1, pp. 404–406, 2016. doi: 10.1109/SCM.2016.7519794.

[28] R. E. Caraka, H. Yasin, and A. W. Basyiruddin, "Peramalan crude palm oil (CPO) menggunakan support vector regression kernel radial basis," *J. Mat.*, vol. 7, no. 1, p. 43, 2017. doi: 10.24843/jmat.2017.v07.i01.p81.

[29] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021. doi: 10.7717/PEERJ-CS.623.

[30] R. Fitriana, D. Sugiarto, M. K. Hidayat, and Y. E. H. Turnip, "Perancangan model sistem intelijensia bisnis," *Wawasan Ilmu*, 2024.

[31] K. V. Metre, A. Mathur, R. P. Dahake, Y. Bhapkar, J. Ghadge, P. Jain, and S. Gore, "An introduction to Power BI for data analysis," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 1s, pp. 142–147, 2023.

[32] R. Fitriana, J. Saragih, and N. Luthfiana, "Model business intelligence system design of quality products by using data mining in R Bakery Company," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 277, no. 1, 2017. doi: 10.1088/1757-899X/277/1/012005.

[33] S. Widjaja and T. Mauritsius, "The development of performance dashboard visualization with Power BI as platform," *Int. J. Mech. Eng. Technol.*, vol. 10, no. 5, pp. 235–249, 2019. [Online]. Available: http://www.iaeme.com/IJMET/issues.asp?JType=IJMET&VType=10&IType=5.

[34] R. Adinugroho, *Perbandingan rasio split data training dan data testing menggunakan metode LSTM dalam memprediksi harga indeks saham Asia*. 2022. [Online]. Available: https://repository.uinjkt.ac.id/dspace/handle/123456789/67314.

[35] A. N. Safira, B. Warsito, and A. Rusgiyono, "Analisis support vector regression (SVR) dengan algoritma grid search time series cross validation untuk prediksi jumlah kasus terkonfirmasi COVID-19 di Indonesia," *J. Gaussian*, vol. 11, no. 4, pp. 512–521, 2023. doi: 10.14710/j.gauss.11.4.512-521.

[36] F. Kaliyadan and V. Kulkarni, "Types of variables, descriptive statistics, and sample size," *Indian Dermatol. Online J.*, vol. 10, no. 4, pp. 481–485, 2017. doi: 10.4103/idoj.IDOJ.

[37] Y. Dong, "Descriptive statistics and its applications," *Highlights Sci. Eng. Technol.*, vol. 47, pp. 16–23, 2023. doi: 10.54097/hset.v47i.8159.

[38] A. Zakrani, A. Najm, and A. Marzak, "Support vector regression based on grid-search method for agile software effort prediction," *Colloq. Inf. Sci. Technol. CIST*, vol. 2018-Octob, pp. 492–497, 2018. doi: 10.1109/CIST.2018.8596370.

[39] A. W. Ishlah, S. Sudarno, and P. Kartikasari, "Implementasi gridSearchCV pada support vector regression (SVR) untuk peramalan harga saham," *J. Gaussian*, vol. 12, no. 2, pp. 276–286, 2023. doi: 10.14710/j.gauss.12.2.276-286.

AUTHORS BIOGRAPHY

Tryantomo Lokhilmahful Palgunadi is an Industrial Engineering graduate from Universitas Trisakti, Jakarta. He obtained his Bachelor degree in the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Trisakti Jakarta, in 2024. He was a lecturer assistant at the Quality Engineering Laboratory.

Rina Fitriana is an Associate professor and Head of Department in the Department of Industrial Engineering, Universitas Trisakti, Indonesia. Her educational background includes a Bachelor of Industrial Engineering from

Universitas Trisakti, a Magister of Management from PPM School of Management, and a Doctor of Agricultural Industry Technology from IPB University. She has many researchs in Industrial Engineering. Her research interests include Quality Engineering, Analyze System analysis of information system design, Data Mining, and Business Intelligence. Rina can be contacted at email: rinaf@trisakti.ac.id

Anik Nur Habyba is a lecturer and Head of the Quality Engineering Laboratory in the Department of Industrial Engineering, Universitas Trisakti, Indonesia. Her educational background is a Bachelor of Agricultural Industry Technology from Brawijaya University and a Master of Agricultural Industry Technology from IPB University. She has more than five years of teaching/ research in Industrial Engineering. Her research interests include Statistics, Quality Engineering, Design of Experiments, and Technology-based entrepreneurship. Anik can be contacted at email: anik@trisakti.ac.id

Yun-Chia Liang is a Professor and Chair of the Department of Industrial Engineering & Management, Yuan-Ze University, Taiwan, ROC. His Research & Interests are Meta-Heuristics, Intelligent Computing, Production Scheduling, Logistics Management, and Artificial Neural Network Applications.

DOI:10.25077/josi.v23.n2.p149-166.2024