Journal of Soft Computing Exploration
Homepage: https://shmpublisher.com/index.php/joscex
p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Particle swarm optimization-based support vector regression for unemployment rate prediction using panel data

Muhtajuddin Danny[1], Asep Muhidin[2]
[1,2]Department of Informatics Engineering, Universitas Pelita Bangsa, Indonesia

Article Info

Article history:

Received May 11, 2026
Revised June 07, 2026
Accepted July 04, 2026

Keywords:

Open unemployment rate
Support vector regression
Particle swarm optimization
Panel data
Machine learning

ABSTRACT

Predicting the Open Unemployment Rate (OUR) is important for supporting data-driven employment policies, particularly in regions with complex economic and social conditions. This study aims to develop a predictive model for OUR using panel data from districts/cities in West Java Province during the 2018–2025 period by applying Support Vector Regression (SVR) optimized with Particle Swarm Optimization (PSO). The dataset includes economic, social, and demographic variables, namely labor force participation rate, average years of schooling, population, minimum wage, Human Development Index, GRDP per capita, poverty rate, and population density. The proposed approach combines SVR as a nonlinear regression technique with PSO for hyperparameter optimization to improve prediction accuracy. The experimental results show that the model achieved a Mean Squared Error (MSE) of 1.4888 and a coefficient of determination (R^2) of 0.4343, indicating moderate predictive performance. In addition, the optimization process demonstrated a stable reduction in RMSE values during iterations, confirming the effectiveness of PSO in enhancing the SVR model. The findings suggest that the SVR–PSO model is capable of capturing general unemployment patterns in panel data and can support adaptive, data-driven employment policy analysis.

Corresponding Author:

Muhtajuddin Danny,
Department of Informatics Engineering,
Universitas Pelita Bangsa,
Kampus B Fakultas Vokasi, Jl. Dharmawangsa Dalam, 28-30, Surabaya, 60286, Indonesia.
Email: utat@pelitabangsa.ac.id
https://doi.org/10.52465/joscex.v7i3.113

1. INTRODUCTION

Unemployment remains one of the most important indicators for evaluating regional economic performance and social stability because it reflects the ability of the economy to absorb labor effectively [1], [2]. A high Open Unemployment Rate (OUR) contributes to poverty, inequality, and reduced public welfare, while also creating long-term economic and social pressures [2], [3]. In West Java Province, unemployment dynamics are highly complex due to significant differences among districts and cities in terms of economic development, education, population density, and labor market conditions [4], [5]. These conditions indicate that unemployment is influenced by multidimensional economic, social, and demographic factors with nonlinear interactions [6]. Therefore, a predictive approach capable of capturing complex relationships and temporal-spatial variations is needed to support more adaptive and evidence-based employment policy formulation [7].

Previous studies on unemployment forecasting have mainly employed statistical and econometric approaches such as linear regression, panel regression, and ARIMA models due to their simplicity and interpretability [8], [9]. However, these methods often face limitations in capturing nonlinear relationships, multidimensional interactions, and structural changes in economic and labor market data [10], [11]. Recent advances in machine learning, including Support Vector Machine (SVM), Random Forest, Artificial Neural Networks, and deep learning models, have demonstrated superior capability in handling complex and high-dimensional datasets [12], [13]. Among these approaches, Support Vector Regression (SVR) has attracted considerable attention because of its robustness and effectiveness in modeling nonlinear relationships with minimal overfitting risk [14]. Nevertheless, SVR performance strongly depends on optimal hyperparameter selection. To address this issue, researchers have integrated SVR with Particle Swarm Optimization (PSO), which effectively optimizes model parameters through swarm-based search mechanisms [15]. Despite these promising developments, studies applying hybrid SVR–PSO models for unemployment prediction using regional panel data in Indonesia, particularly at the district/city level, remain limited [16].

Although machine learning approaches have demonstrated stronger predictive capability than conventional econometric methods, unemployment prediction studies in Indonesia remain limited, particularly at the regional panel-data level [17], [18]. Most previous studies still focus on statistical inference using linear econometric techniques and rarely integrate economic, social, and demographic variables within a unified predictive framework [19]. Consequently, the complex and nonlinear dynamics of unemployment across districts and cities have not been fully captured. In addition, although Support Vector Regression (SVR) and Particle Swarm Optimization (PSO) have shown promising results in various forecasting applications [20], [21]. their implementation for unemployment prediction using regional panel data in Indonesia, especially in West Java Province, remains scarce. Therefore, a more adaptive hybrid predictive framework is needed to improve prediction accuracy and model nonlinear unemployment dynamics across spatial and temporal dimensions.

This study proposes a hybrid machine learning framework by integrating Support Vector Regression (SVR) and Particle Swarm Optimization (PSO) to predict the Open Unemployment Rate (OUR) using panel data from 27 districts/cities in West Java Province during 2018–2025 [22]. The dataset includes economic, demographic, and social variables, such as labor force participation rate, average years of schooling, population, minimum wage, Human Development Index (HDI), GRDP per capita, poverty rate, and population density [23]. Data preprocessing involves cleaning, normalization, outlier handling, and transformation to ensure data quality and model compatibility. SVR is employed to capture nonlinear relationships and multidimensional patterns [24], while PSO is used to optimize SVR hyperparameters to improve prediction accuracy and model stability [25]. The model performance is evaluated using MAE, MSE, RMSE, and coefficient of determination (R^2) [26].

Unemployment remains a critical socioeconomic issue because it reflects the ability of a region to absorb available labor resources and maintain sustainable economic growth. Accurate unemployment forecasting is therefore essential for supporting labor market policies and evidence-based decision making. Several researchers have focused on unemployment forecasting using conventional econometric approaches or standalone machine learning models [27], [28]. However, there are limited studies related to the implementation of hybrid SVR–PSO models for predicting the Open Unemployment Rate using district/city-level panel data in Indonesia [29]. Therefore, a significant research gap remains in developing an adaptive predictive framework capable of simultaneously capturing nonlinear, multidimensional, spatial, and temporal unemployment patterns. The novelty of this study lies in the integration of Particle Swarm Optimization-based hyperparameter tuning with Support Vector Regression using multidimensional panel data from 27 districts/cities in West Java Province during 2018–2025. Unlike previous studies that primarily relied on conventional econometric methods or standalone machine learning algorithms, most existing approaches have limited capability in simultaneously capturing nonlinear relationships and the spatial-temporal characteristics of regional unemployment dynamics [30]. Therefore, this research intends to develop a hybrid Support Vector Regression–Particle Swarm Optimization (SVR–PSO) model for predicting the Open Unemployment Rate using multidimensional regional panel data. The objectives of this research are to develop and evaluate the proposed SVR–PSO model and to assess its predictive performance using MAE, MSE, RMSE, and R^2 evaluation metrics [31], [32].

2. METHOD

This study applies a structured predictive modeling framework to estimate the Open Unemployment Rate using panel data. The process includes dataset preparation, data preprocessing (cleaning, normalization, and transformation), and model development using Support Vector Regression (SVR) to capture nonlinear relationships. Parameter optimization is performed using Particle Swarm Optimization (PSO) to obtain optimal hyperparameters. The model is then evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2) to assess prediction accuracy and generalization performance.

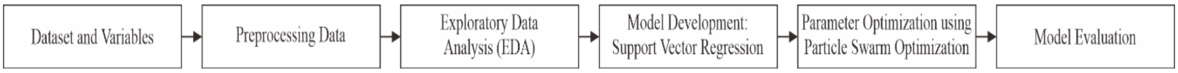

Figure 1. Research design and workflow

Figure 1 is a research workflow that illustrates dataset preparation, pre-processing, SVR modeling, PSO optimization, and model evaluation.

Dataset and Variables

The dataset used in this study is secondary panel data obtained from Statistics Indonesia (BPS), covering 27 districts/cities in West Java Province for the period 2018–2025, resulting in 216 observations. The dependent variable is the Open Unemployment Rate (OUR), while the independent variables include labor force participation rate, average years of schooling, population, minimum wage, Human Development Index, GRDP per capita, poverty rate, and population density.

The panel data structure enables the model to capture both cross-sectional and time-series variations simultaneously, providing a more comprehensive representation of unemployment dynamics.

Table 1. Research variables and operational definitions

No	Variable	Type	Unit	Operational Definition
1	Open Unemployment Rate (OUR)	Dependent	%	Percentage of the labor force that is unemployed relative to the total labor force
2	Labor Force Participation Rate (LFPR)	Independent	%	Percentage of working-age population that is economically active
3	Average Years of Schooling	Independent	Years	Average number of years of formal education attained by the population
4	Population	Independent	Thousand people	Total population in each district/city
5	Minimum Wage (District/City)	Independent	Rupiah	Regional minimum wage standard
6	Human Development Index (HDI)	Independent	Index	Indicator of the quality of human development
7	Gross Regional Domestic Product (GRDP) per Capita	Independent	Rupiah	Economic output value per individual
8	Poverty Rate	Independent	%	Percentage of population living below the poverty line
9	Population Density	Independent	People/km²	Number of people per unit area

Table 1 presents the research variables used in the Open Unemployment Rate (OUR) prediction model, which consists of 1 dependent variable and 8 independent variables.

Table 2. Panel data structure

Dimension	Description
Cross-section	27 districts/cities in West Java
Time series	Year 2018–2025
Number of observations	27 × 8 = 216 data points
Data type	Secondary data
Data source	Statistics Indonesia (BPS) and official publications

Table 2 illustrates the panel data structure used in the study, which consists of a cross-section dimension in the form of 27 districts/cities in West Java Province and a time series dimension for the period 2018 to 2025. With the combination of these two dimensions, the total observations used in this study amount to 216 data.

Preprocessing Data

Data preprocessing was conducted to prepare the dataset for machine learning modeling through data cleaning, feature normalization, data transformation, and data splitting. Data cleaning was performed to address missing values, inconsistencies, and potential outliers. Feature normalization was carried out using the Min–Max Scaling technique, which transforms variables into a comparable range prior to model training and is commonly applied in machine learning algorithms sensitive to feature scales, including Support Vector Regression (SVR) [33].

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where x' denotes the normalized value, x is the original value, and xmin and xmax represent the minimum and maximum values of the corresponding feature, respectively. Furthermore, the panel dataset was transformed into a machine-learning-ready format by organizing the predictor variables into feature matrices and the response variable into target vectors for model development. Subsequently, the dataset was divided into training (80%) and testing (20%) subsets using the hold-out validation approach, a commonly adopted strategy for evaluating model generalization performance on unseen data [34].

Table 3. Data preprocessing steps

No	Step	Description
1	Data Cleaning	Missing values, inconsistencies, and duplicate records were identified and corrected prior to analysis.
2	Normalization	Numerical variables were normalized using the Min–Max Scaling method to transform values into a comparable range before model training [35].
3	Outlier Detection	Outliers were identified through exploratory data analysis and evaluated prior to model development.
4	Data Transformation	The panel dataset was transformed into feature matrices and target vectors suitable for SVR modeling.
5	Data Splitting	The dataset was divided into training (80%) and testing (20%) subsets using the hold-out validation approach [36].

Table 3 describes the data preprocessing steps performed to improve the quality and readiness of the dataset before the modeling process. These steps include data cleaning, normalization, outlier detection, and data transformation, aimed at producing clean, structured, and uniformly scaled data, thereby improving the performance of the predictive model.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to examine the statistical characteristics and relationships among variables prior to model development [37]. In this study, EDA was performed through descriptive statistical analysis and correlation analysis to identify potential relationships among economic, social, and demographic variables associated with the Open Unemployment Rate. The results revealed varying correlation patterns among the predictor variables, indicating the presence of complex interactions that may not be adequately captured by conventional linear approaches. These findings support the use of Support Vector Regression (SVR), which is capable of modeling nonlinear relationships in multidimensional datasets [38].

Table 4. Summary of exploratory data analysis results

Variable	Mean	Std. Dev	Min	Max	Key Insight
Open Unemployment Rate	7.85	2.15	2.10	14.20	Moderately distributed without extreme skewness
LFPR	66.30	4.80	55.20	75.10	Stable variation across regions
Education	8.45	1.20	6.10	11.30	Strong relationship with HDI
Population	1450	820	320	3200	High variability across districts
Minimum Wage	2.8M	0.7M	1.8M	4.5M	Moderate economic variation
HDI	71.50	4.30	62.10	81.20	Strong indicator of human development
GRDP	38.20	15.60	12.40	85.30	Reflects economic disparities
Poverty	8.90	3.10	2.80	15.60	Negatively related to HDI
Density	4200	3100	500	14000	Highly varied spatial distribution

Table 4 presents the summary statistics of the study variables. Several variables show high regional variability, indicating heterogeneous characteristics and complex nonlinear relationships that support the use of Support Vector Regression.

Model Development: Support Vector Regression

The predictive model used in this study is Support Vector Regression (SVR), which is designed to model nonlinear relationships between input variables and the target variable [39]. The model operates by mapping input data into a higher-dimensional feature space using a nonlinear transformation function, allowing complex patterns in panel data to be effectively captured [40].

The regression function of SVR is defined as follows [41]:

$$f(x) = w^T \phi(x) + b \tag{2}$$

where f(x)represents the predicted value of the Open Unemployment Rate, wis the weight vector, ϕ(x)is the nonlinear mapping function, and b is the bias term. The model aims to minimize prediction error within an ε-insensitive loss margin, allowing small deviations without penalty.

In this study, the Radial Basis Function (RBF) kernel is employed due to its flexibility in capturing nonlinear relationships in multidimensional data [42]:

$$K(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2) \tag{3}$$

where $\gamma$is the kernel parameter controlling the influence of individual data points. The SVR model uses three main hyperparameters: the regularization parameter C, the kernel parameter γ, and the ε-insensitive loss parameter ϵ. These parameters significantly influence model performance and must be carefully determined.

The SVR model uses normalized features as input and the Open Unemployment Rate as output. The dataset is divided into 80% training data and 20% testing data using the hold-out method to evaluate model performance and reduce overfitting [43].

Parameter Optimization using Particle Swarm Optimization

To improve SVR performance, Particle Swarm Optimization (PSO) is used to optimize the hyperparameters (C, γ, and ϵ) through an iterative population-based search process [44].

The velocity update is defined as follows [45]:

$$v_i^{t+1} = w \cdot v_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t) \tag{4}$$

The position update is defined as follows [46]:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{5}$$

where $v_i$represents the velocity of particle i, $x_i$is its position, $w$is the inertia weight controlling exploration and exploitation, $c_1$and $c_2$are acceleration coefficients, and $r_1$and $r_2$are random numbers in the interval [0,1].

The objective function used in the optimization process is Mean Squared Error (MSE), which is defined as follows [47]:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{6}$$

where $y_i$is the actual value and $\hat{y}_i$is the predicted value. The optimization process aims to minimize the MSE value to obtain the best-performing SVR parameters.

Model Evaluation

Model evaluation was conducted using the testing dataset (20% of total data) obtained through the hold-out method to assess predictive performance and generalization capability. The model was evaluated using MAE, MSE, RMSE, and R² metrics.

The Mean Absolute Error (MAE) is defined as follows [48]:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \mid y_i - \hat{y}_i \mid \tag{7}$$

where $y_i$represents the actual value and $\hat{y}_i$represents the predicted value. MAE measures the average magnitude of prediction errors without considering their direction.

The Mean Squared Error (MSE) is defined as follows [49]:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{8}$$

The Root Mean Squared Error (RMSE) is defined as follows [50]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{9}$$

.

RMSE provides an interpretable measure of prediction error in the same unit as the target variable, allowing easier comparison with actual values.

The coefficient of determination (R^2) is defined as follows [51]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{10}$$

where $\bar{y}$is the mean of the actual values. The R^2metric indicates the proportion of variance in the dependent variable that can be explained by the model.

Table 5. Model evaluation metrics

Metric	Description
MAE	Measures the average absolute prediction error
MSE	Measures the average squared prediction error
RMSE	Measures the square root of MSE in the same unit as the target variable
R^2	Measures the proportion of variance explained by the model

Table 5 presents the evaluation metrics used in this study, where lower MAE, MSE, and RMSE values indicate better prediction performance, while a higher R^2 value indicates stronger explanatory capability.

3. RESULTS AND DISCUSSIONS

The SVR model optimized with PSO achieved an MSE of 1.4888 and an R^2 of 0.4343, indicating moderate predictive performance. The results show that the model is capable of minimizing prediction errors and capturing the general pattern of unemployment dynamics across districts and cities.

Table 6. Model performance results

Metric	Value
MAE	0.87
MSE	1.4888
RMSE	1.22
R^2	0.4343

Table 6 shows that the proposed model achieves a relatively low error rate, indicating acceptable prediction accuracy. The RMSE value of 1.22 reflects a reasonable deviation between predicted and actual values, while the R^2 value indicates moderate explanatory capability, suggesting that other factors may still influence unemployment variability.

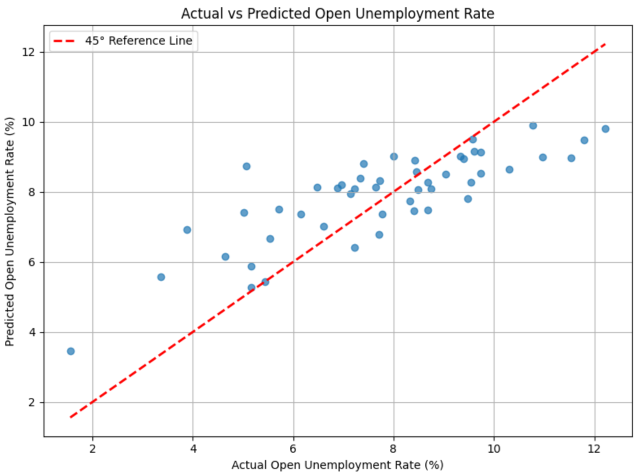

Figure 2. Actual versus predicted open unemployment rate (OUR)

Figure 2 shows the comparison between actual and predicted values of the Open Unemployment Rate. The model is able to follow the general trend of the data, although some deviations are observed in several periods. This indicates that the model has moderate predictive capability and is able to capture the overall pattern of unemployment dynamics.

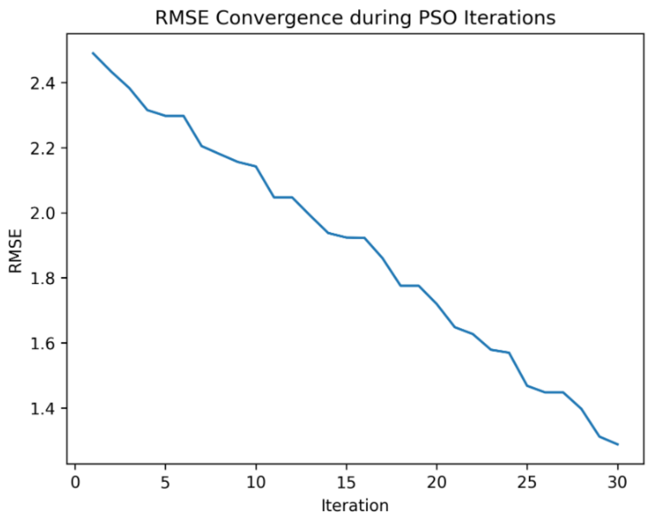

Figure 3. PSO convergence

Figure 3 illustrates the convergence of RMSE during the PSO optimization process. The decreasing trend indicates that the algorithm effectively reduces prediction error and converges toward optimal SVR parameters. Minor fluctuations in early iterations reflect exploration, while later stability indicates convergence.

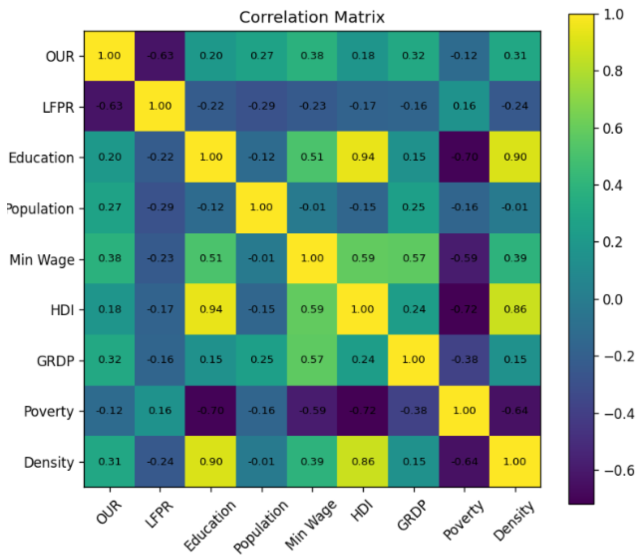

Figure 4. Correlation matrix

Figure 4 presents the correlation matrix of the research variables. The results show both strong and weak relationships among variables, indicating complex interactions. Strong positive correlations are observed between education and HDI, while poverty shows negative relationships with development indicators. This variability suggests nonlinear relationships, supporting the use of SVR for modeling.

4. CONCLUSION

This study successfully achieved its objective of developing and evaluating a hybrid Support Vector Regression–Particle Swarm Optimization (SVR–PSO) model for predicting the Open Unemployment Rate using multidimensional panel data from districts and cities in West Java Province during 2018–2025. The results presented in the Results and Discussion section demonstrate that the proposed model is capable of capturing nonlinear relationships among socioeconomic variables and generating reliable unemployment forecasts. The model achieved a Mean Squared Error (MSE) of 1.4888 and a coefficient of determination (R^2) of 0.4343, indicating moderate predictive performance in explaining regional unemployment dynamics.

The findings show that the integration of Particle Swarm Optimization (PSO) contributes to improving SVR performance through the optimization of key hyperparameters, resulting in a more adaptive predictive framework for modeling complex unemployment patterns. These results suggest that hybrid machine learning approaches offer advantages over conventional methods in handling multidimensional and nonlinear relationships within regional socioeconomic data. Therefore, the proposed SVR–PSO model provides a promising alternative for unemployment forecasting and regional labor market analysis.

From a theoretical perspective, this study contributes to the growing application of hybrid machine learning and intelligent optimization techniques in socioeconomic forecasting. Practically, the proposed framework can support policymakers in monitoring unemployment trends and formulating evidence-based employment policies. However, the moderate explanatory power of the model indicates that unemployment is influenced by additional factors beyond those considered in this study. Therefore, future research is encouraged to incorporate more comprehensive variables, broader datasets, and advanced optimization or machine learning techniques. Furthermore, the proposed framework has the potential to be integrated into regional labor market monitoring systems and decision-support platforms for more effective policy evaluation and planning.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation to Universitas Pelita Bangsa for its institutional support and the academic environment that facilitated the conduct of this research. This study was supported by internal research funding from Universitas Pelita Bangsa, for which we gratefully acknowledge. We would also like to thank the Central Bureau of Statistics (BPS) for providing reliable and accessible data that made this research possible. Furthermore, we would like to express our gratitude to all parties who have contributed directly or indirectly, including colleagues and academic colleagues, for their constructive feedback and support in improving the quality of this work.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Muhtajuddin Danny: Conceptualization, data curation, software, formal analysis, methodology, visualization, and writing–original draft. **Asep Muhidin**: Supervision, validation, methodology, and writing–review & editing.

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

The data used in this study were obtained from reports published by the Central Statistics Agency (BPS) and other official publicly available publications related to districts/cities in West Java Province during the 2018–2025 period. The processed dataset, along with other supporting research materials, is available from the corresponding author upon reasonable request.

REFERENCES

[1] N. E. Ramadhani and P. D. Karjati, “Impact of Wages, Unemployment Rates, and Inflation on Labor Forece Participation,” *Indep. J. Econ.*, vol. 5, no. 2, pp. 13–19, 2025.

[2] M. N. Agustiani, S. Adilia, E. Della Rosa, and M. Panorama, “The Effect of Minimum Wages, Unemployment Rate, and Inflation on Poverty In Indonesia,” *J. Manajemen, Akuntansi, Ekon.*, vol. 1, no. 3, pp. 39–44, 2022.

[3] N. Azmy and R. I. S. Setiawati, “Analysis of the Effect of Population, Open Unemployment Rate, and Economic Growth on Poverty Rates in Lamongan Regency,” *J. Bus. Manag. Econ. Dev.*, vol. 3, no. 01, pp. 65–75, 2025.

[4] M. A. Habibi, M. Taufiq, and A. F. Utami, “Analysis of the Influence of Economic Growth, Average Length of Schooling, and Provincial Minimum Wage on the Open Unemployment Rate (TPT) in 6 Provinces of Java,” *J. Pemberdaya. Ekon. dan Masy.*, vol. 1, no. 3, p. 9, 2024.

[5] S. E. Juliantika and S. Arka, “Analysis of the Influence of Economic Growth, Unemployment Rates, Poverty Levels on the Level of Community Welfare in the Regency/City of Bali Province,” *Int. J. Econ. Manag. Sci.*, vol. 1, no. 3, pp. 382–408, 2024.

[6] N. F. Campos, V. Z. Eichenauer, and J.-E. Sturm, “Are unemployment differentials among advanced economies still explained by the Shocking Institutions Hypothesis?,” *Econ. Syst.*, vol. 49, no. 4, p. 101325, 2025.

[7] C. Jo, D. H. Kim, and J. W. Lee, “Forecasting unemployment and employment: A system dynamics approach,” *Technol. Forecast. Soc. Change*, vol. 194, p. 122715, 2023.

[8] S. Baya, A. Fadlallah, H. El Baraka, K. Bourouis, and M. Ezzraouli, “Modeling and Forecasting the Real Effective Exchange Rate in Morocco: A Comparative Analysis by ARIMA, Random Forest and the Dynamic Factor Model,” in *ICATH 2025*, 2025, p. 53.

[9] T. Annastasya, R. Passarella, and Z. Yamani, “Unemployment rate forecasting in Indonesia using macroeconomic indicators with a machine learning approach,” *Discov. Anal.*, vol. 3, no. 1, p. 15, 2025.

[10] M. F. Fadhilah, M. Salsabila, F. S. Aufa, O. Yulianti, and R. S. A. Sabila, “Analysis of the Effect of Average Years of Schooling and Provincial Minimum Wages on the Human Development Index Across 38 Provinces in Indonesia (2024),” *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 4, pp. 6181–6191, 2025.

[11] E. Arianto and H. Cahyono, "The Effect of Human Development Index and Unemployment on Poverty in Indonesia 2018-2023," *Financ. Int. J. Manag. Financ.*, vol. 2, no. 4, pp. 59–64, 2025.

[12] A. Utomo, D. Munthe, and D. Handoko, "Advancing macroeconomic forecasting by integrating big data and machine learning," *ECOBIZTECH J. Econ. Business, Technol.*, vol. 1, no. 1, pp. 21–35, 2026.

[13] M. Simionescu, "Machine Learning vs. Econometric Models to Forecast Inflation Rate in Romania? The Role of Sentiment Analysis," *Mathematics*, vol. 13, no. 1, p. 168, 2025.

[14] M. Ali and Y. Fissha, "Propose adjustable Support Vector Machine approach for classifying imbalanced work travel mode choice data," *Transp. Res. Interdiscip. Perspect.*, vol. 35, p. 101786, 2026.

[15] A. et al. Casas-Ordaz, "Particle swarm optimization: A survey of innovations over the last 10 years," *Comput. Sci. Rev.*, vol. 60, p. 100910, 2026.

[16] A. G. di Stefano, M. Ruta, G. Masera, and S. Hoque, "Leveraging Machine Learning to Forecast Neighborhood Energy Use in Early Design Stages: A Preliminary Application," *Buildings*, vol. 14, no. 12, p. 3866, 2024.

[17] R. Passarella, M. I. Setiawan, and Z. Yamani, "Comparative Analysis of Machine Learning Models for Predicting Indonesia's GDP Growth," *Acadlore Trans. AI Mach. Learn.*, vol. 4, no. 3, pp. 157–173, 2025.

[18] L. Popescu, M. Gaman, L. S. Mihai, C. O. Dragan, D. Militaru, and I. Buligiu, "Econometric Analysis and Forecasts on Exports of Emerging Economies from Central and Eastern Europe," *Econometrics*, vol. 14, no. 1, p. 9, 2026.

[19] R. Snaiki and A. Merabtine, "Recent advances on machine learning techniques for urban heat island applications: a review and new horizons," *Sustain. Cities Soc.*, vol. 134, p. 106943, 2025.

[20] M. M. F. Dip, S. K. Adhikary, S. P. Shuvo, M. J. P. Ratul, and U. Akter, "A Hybrid PSO-SVM Framework Incorporating Advanced Preprocessing Techniques for Enhanced Prediction of River Water Level," 2026, pp. 334–346.

[21] R. V Poursorkhabi, A. R. Fard, and A. R. Fard, "Designing a model for earthquake timing and magnitude prediction based on neural networks and particle swarm optimization (PSO) algorithm," *Discov. Geosci.*, vol. 4, no. 1, p. 42, 2026.

[22] X. Su, H. Jiang, T. Qin, and G. Lin, "Particle Swarm Optimization–Support Vector Regression (PSO-SVR)-Based Rapid Prediction Method for Radiant Heat Transfer for a Spacecraft Vacuum Thermal Test," *Appl. Sci.*, vol. 14, no. 20, p. 9407, 2024.

[23] A. Wibowo and E. Setyowati, "Analysis of the Effect of Population, Unemployment Rate, Minimum Wage, and Human Development Index on Poverty Levels in Yogyakarta DIY Province in 2017--2021," 2024, pp. 217–227.

[24] Q. Wang, "Support Vector Machine Algorithm in Machine Learning," in *2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2022, pp. 750–756.

[25] Q. Zhang, J. Sheng, R. Zhou, Q. Zhang, B. Wang, and R. Zhang, "Enhanced particle swarm optimization for feature selection in SVM-based Alzheimer's disease diagnosis," *Sci. Rep.*, vol. 15, no. 1, p. 24243, 2025.

[26] Y. Altork, "Comparative analysis of machine learning models for wind speed forecasting: Support vector machines, fine tree, and linear regression approaches," *Int. J. Thermofluids*, vol. 27, p. 101217, May 2025.

[27] H. Tjahjanto, T. Tuhana, I. Mafruhah, N. Istiqomah, and D. Ismoyowati, "High unemployment, disrupted economic growth and sustainable development goals: Analyzing unemployment reduction," *Econ. Sociol.*, vol. 16, no. 1, pp. 106–120, 2023.

[28] A. Syarifudin, H. Hamidah, N. Zainal, S. Anwar, N. A. Shafie, and A. M. Liton, "Graduate Unemployment Hinders Economic Growth in Indonesia: Does Government Spending on Education Matters?," *Int. J. Econ. Financ. Issues*, vol. 15, no. 6, pp. 472–480, Oct. 2025.

[29] M. Yollanda and D. Devianto, "Hybrid Model of Seasonal ARIMA-ANN to Forecast Tourist Arrivals through Minangkabau International Airport," in *Proceedings of the 1st International Conference on Statistics and Analytics (ICSA 2019)*, 2020.

[30] O. G. Erbulut and Z. Colak, "A hybrid machine learning approach for housing price prediction: the stacking regressor method," *Int. J. Hous. Mark. Anal.*, Mar. 2025.

[31] A. Jovanovic and others, "Particle swarm optimization tuned multi-headed long short-term memory networks approach for fuel prices forecasting," *J. Netw. Comput. Appl.*, vol. 233, p. 104048, Jan. 2025.

[32] Y. Zhao, W. Zhang, and X. Liu, "Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting," *Appl. Soft Comput.*, vol. 154, p. 111362, Mar. 2024.

[33] P. Koukaras and C. Tjortjis, "Data Preprocessing and Feature Engineering for Data Mining: Techniques, Tools, and Best Practices," *AI*, vol. 6, no. 10, p. 257, Oct. 2025.

[34] K. M. M. Uddin, A. Chowdhury, M. M. R. Druvo, M. T. Jaima, M. T. A. Bhuiyan, and M. M. Islam, "Improving diagnostic accuracy for PCOS: A hybrid machine learning architecture with feature selection, data balancing, and explainable AI techniques," *Results Control Optim.*, vol. 22, p. 100647, Mar. 2026.

[35] Z. Wang and others, "Predictive modeling of compressive strength in tailings concrete using explainable machine learning approaches," *Results Eng.*, vol. 27, p. 105516, Sep. 2025.

[36] A. Andrews, N. Islam, G. Bcharah, H. Bcharah, and M. Pangasa, "Development and Internal Validation of Machine Learning Algorithms to Predict 30-Day Readmission in Patients Undergoing a C-Section: A Nation-Wide Analysis," *J. Pers. Med.*, vol. 15, no. 10, p. 476, Oct. 2025.

[37] N. T. M. Sagala and F. Y. Aryatama, "Exploratory Data Analysis (EDA): A Study of Olympic Medallist," *SISTEMASI*, vol. 11, no. 3, p. 578, Sep. 2022.

[38] Y. Liu, L. Wang, and K. Gu, "A support vector regression (SVR)-based method for dynamic load identification using heterogeneous responses under interval uncertainties," *Appl. Soft Comput.*, vol. 110, p. 107599, Oct. 2021.

[39] Z. Qi, Y. Feng, S. Wang, and C. Li, "Enhancing hydropower generation Predictions: A comprehensive study of XGBoost and Support Vector Regression models with advanced optimization techniques," *Ain Shams Eng. J.*, vol. 16, no. 1, p. 103206, Jan. 2025.

[40] D. Koldasbayeva, P. Tregubova, M. Gasanov, A. Zaytsev, A. Petrovskaia, and E. Burnaev, "Challenges in data-driven geospatial modeling for environmental research and practice," *Nat. Commun.*, vol. 15, no. 1, p. 10700, Dec. 2024.

[41] A. M. Priyatno and Y. Ningsih, "Evaluating Imputation Approaches and Support Vector Regression Parameters in Weather Forecasting," *J. Eng. Sci. Appl.*, vol. 2, no. 2, pp. 21–28, Oct. 2025.

[42] J. Zatarain Salazar, J. H. Kwakkel, and M. Witvliet, "Evaluating the choice of radial basis functions in multiobjective optimal control applications," *Environ. Model. Softw.*, vol. 171, p. 105889, Jan. 2024.

[43] B. Bargam, A. Boudhar, C. Kinnard, H. Bouamri, K. Nifa, and A. Chehbouni, "Evaluation of the support vector regression

(SVR) and the random forest (RF) models accuracy for streamflow prediction under a data-scarce basin in Morocco," *Discov. Appl. Sci.*, vol. 6, no. 6, p. 306, Jun. 2024.
[44] R. S. Ajin, S. Segoni, and R. Fanti, "Optimization of SVR and CatBoost models using metaheuristic algorithms to assess landslide susceptibility," *Sci. Rep.*, vol. 14, no. 1, p. 24851, Oct. 2024.
[45] Y. Y. Akan and J. M. Herrmann, "Preselection and diversity in PSO," *OPSEARCH*, Oct. 2025.
[46] S. Tijjani, M. N. Ab Wahab, and M. H. Mohd Noor, "An enhanced particle swarm optimization with position update for optimal feature selection," *Expert Syst. Appl.*, vol. 247, p. 123337, Aug. 2024.
[47] H. Nuha, "Mean Squared Error (MSE) dan Penggunaannya," *J. Pemanfaat. Teknol. untuk Masy. J. Pengabdi. Masy.*, vol. 3, no. 4, Oct. 2025.
[48] S. M. Robeson and C. J. Willmott, "Decomposition of the mean absolute error (MAE) into systematic and unsystematic components," *PLoS One*, vol. 18, no. 2, p. e0279774, Feb. 2023.
[49] T. O. Hodson, T. M. Over, and S. S. Foks, "Mean Squared Error, Deconstructed," *J. Adv. Model. Earth Syst.*, vol. 13, no. 12, Dec. 2021.
[50] D. S. K. Karunasingha, "Root mean square error or mean absolute error? Use their ratio as well," *Inf. Sci. (Ny).*, vol. 585, pp. 609–629, Mar. 2022.
[51] G. S. Mohammad, "Estimate Accuracy of the Sample Determination Coefficient R2 and R2 for the Multiple Regression Models," *J. Southwest Jiaotong Univ.*, vol. 55, no. 1, 2020.