Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 16, 2023, page 181 – 191

Forecasting Consumer Price Index of Malaysia Using Support Vector Regression

Logashni Morthy, Norhaiza Ahmad*
Department of Mathematical Sciences, Faculty of Science, Universiti Teknologi Malaysia
*Corresponding author: norhaiza@utm.my

Abstract
The CPI measures inflation and predicts consumer spending patterns. Conventional forecasting techniques like ARIMA can be difficult to use because they assume that the data is linear, normal, and stationary. The lack of evidence for these assumptions makes the situation challenging. This thesis suggests employing Support Vector Regression (SVR) as a promising approach to surmount the constraints at hand. SVR can handle non-linear data and has less dependence on input assumptions. This study compares the predictive effectiveness of Support Vector Regression (SVR) using four different kernels: linear, polynomial, radial, and sigmoid. *C* is varied between 10, 100, and 1000. MAE, MAPE, and RMSE are widely used metrics for assessing forecast accuracy. The best cost parameter and kernel for predicting future monthly average Malaysian consumer price index from January 2005 to December 2021 were identified by minimizing MAE, MAPE, and RMSE. The sigmoid kernel outperforms the linear, radial, and polynomial kernels. The sigmoid kernel with a cost parameter of 1000 achieves predictive accuracy on the training dataset with minimal error.

Keywords: Support Vector Regression (SVR), Malaysia's Consumer Price Index, Hyperplane, Kernels

1. Introduction

The Consumer Price Index (CPI) provides insight into the fluctuations in prices for a typical assortment of goods and services that individuals commonly buy, observed during a specific duration. It serves as a methodology to track the fluctuations in prices that consumers face in various areas such as transportation, equipment, and healthcare. This metric, according to Khamis (2020), encompasses a diverse range of items including food, clothing, and entertainment choices. Widely used in assessing economic conditions, the CPI data allows for the evaluation of not only the effectiveness of an economy within the existing regulatory framework but also the standard of living as influenced by employee compensation and their customary spending patterns (Qu, 2022).

The utilization of the consumer price index (CPI) in Malaysia extends beyond measuring inflation rates. Additionally, it functions as a mechanism for assessing the economic strategies implemented by the Malaysian government. The CPI provides a comprehensive understanding of price fluctuations in the country's economy, benefiting the government, businesses, and consumers (Dongdong, 2010). Research indicates that significant financial events within a specific timeframe impact Malaysia's CPI. Notably, two distinct categories of financial occurrences have affected the Malaysian Consumer Price Index: the Global Financial Crisis of 2008–2009 and the 2020 Recession triggered by the COVID-19 pandemic (Janda, 2019). The global economic crisis of 2008–2009 had a substantial impact on international trade, resulting in a 20% decline in global exports during the first quarter of 2009. Additionally, many manufacturing firms heavily reliant on exports were forced to close (UNDP, 2010).

Support Vector Regression (SVR) is a non-parametric forecasting methodology that aims to minimise error rates by fitting the error within a pre-defined threshold. The Support Vector Regression (SVR) is a machine learning methodology that was introduced in 1992 by Vladimir Vapnik and his associates. The concept underlying this approach was also developed during the same year. The "Support Vector Regression" methodology is a regression technique that is capable of handling both linear and nonlinear regression problems, as suggested by its nomenclature. This project will employ Support Vector Regression to forecast 204 Consumer Price Index (CPI) values in Malaysia spanning from January 2005 to December 2021. The objective of this research is to employ Support Vector Regression to predict Malaysia's consumer price index for the period spanning from January 2005 to December 2021.

2. CPI Dataset of Malaysia

This study utilized data from the Malaysian Department of Statistics (DOSM) regarding the total Consumer Price Index (CPI). The Consumer Price Index (CPI) gauges the fluctuation in average prices for a selection of goods and services that people typically buy within a given timeframe. The dataset used in this research project encompasses the total consumer price index of Malaysia spanning from January 2005 to December 2021, resulting in a total of 204 data observations. Within this dataset, there are twelve main categories, namely food and non-alcoholic beverages, alcoholic beverages and tobacco, clothing and footwear, housing, water, electricity, gas and other fuels, furnishings, household equipment and routine household maintenance, health, transport, communication, recreation services and culture, education, restaurants and hotels, and miscellaneous goods and services. From January 2005 to December 2021, the CPI data of Malaysia shows cyclical movements with the data repeating up and down movements. There have been no variations in this time series pattern since September 2016. This might be due to financial concerns that emerge throughout the time. Based on the research it is found that in the year 2020, Malaysia was facing recession which may be one of the reasons for no variation in the dataset.

The mean of a dataset indicates the average value of Malaysia's consumer price index. Malaysia's overall consumer price index has a mean of 138.7. The median is the mid- number in a Malaysian consumer price index statistic. It is computed by first ordering all of the observations in a dataset and then determining the median value. This dataset's median consumer price index is 119.5. The mean and median estimate the location of the consumer price index of Malaysia dataset's center. The lowest recorded consumer price index was 86.6 in February 2005, while the highest recorded consumer price index was 230.6 in September 2016. The first and third quartiles had values of 104.5 and 186.3, respectively.

The ADF test plays a crucial role in evaluating the stationarity of time series data, effectively determining whether the data exhibits a consistent pattern over time or not. The results of the ADF test of CPI data for Malaysia from January 2005 to December 2016. According to the results, the p-value obtained is 0.01 which is less than $p = 0.05$, thus the null hypothesis is rejected. As a result, overall CPI data remains stationary. Also, based on Chow test, the p-value obtained was 0.009582, thus there is a structural break in the data. It demonstrates that Malaysia's CPI has one breakpoint, which lines up with the break date in September 2016. Indicating that there are factors impacting the consumer price index both before and after the structural break, resulting in data volatility, this demonstrates that the region has a significant breakpoint at $t = 141$.

3. Research Methodology.

The data used in this study comprises total Consumer Price data, which is updated by the Malaysian Department of Statistics (DOSM) each month. This research project used the total consumer price index of Malaysia from January 2005 to December 2021 which yielded a total of 204 data observations The CPI is then predicted using four kernels for comparison, and the support vector regression value is chosen.

3.1 *Support Vector Regression Method*

Support Vector Regression (SVR) is a part of the Support Vector Machine (SVM) that makes predictions. SVR formulates an optimization problem that yields a regression function that translates predictor variables to observed response values. SVR is valuable because it properly balances model complexity and prediction inaccuracy and easily analyses high-dimensional data. (Vapnik et. al, 2002). To solve regression concerns, SVR generates a binary output, while SVM classification produces a real-valued function estimate. As a result, SVR explains optimization in terms of support vectors, with the optimization solution being determined entirely by the number of support vectors, rather than the amount of the input data.

SVR outperforms other regression methods such as multiple and basic linear regression. According to the research conducted by K et al. in 2001, it is argued that Support Vector Regression (SVR) has the potential to effectively address nonlinear regression problems. The accomplishment lies in the transformation of the initial features into a kernel space, thereby facilitating the linear differentiation of the data. This process allows for a distinct and unique analysis, ensuring originality without any plagiarism concerns. Another advantage of SVR is that, unlike conventional data regression techniques, it trains a model to characterize the importance of a variable in determining the link between input and output. When the actual connection between inputs and outputs is not linear, linear assumes that the input data is distributed linearly without creating relevant coefficients or confidence ranges. Unlike other regression methods, Support Vector Regression (SVR) incorporates a confidence interval that effectively illustrates the connection between input variables and output predictions, as highlighted by Glaser, Benjamin, Farhoodi, and Kording in their 2019 study.

Furthermore, SVR comprises four fundamental constituents, namely the support vector, boundary lines, kernel, and hyperplane. A hyperplane, in a higher-dimensional context, represents a linear boundary that separates two distinct classes of data points. This line is used by the SVR to predict the target value. SVR in a higher dimension is used for regression. The kernel is the name given to this functionality. SVR employs a variety of kernels, including the Sigmoidal Kernel, Polynomial Kernel, Gaussian Kernel, and Linear Kernel (Editya et al., 2021). The two lines drawn away from the hyperplane indicate the boundary lines. Its purpose lies in generating a separation amidst data elements. The phrase "support vector" refers to either the vector that was used to construct the hyperplane or the extreme data points in the dataset that helped define it. These data points are almost reaching their limit.

In SVR, training data is referred to as a support vector, which is then tested. Function *f(x)* approximates the target variable with the maximum margin while allowing some deviations.

$$\boldsymbol{f}\,(\boldsymbol{x_i}) \;=\; \boldsymbol{w}^{\boldsymbol{T}}\boldsymbol{\Phi}(\boldsymbol{x_i}) \;+\; \boldsymbol{b} \tag{1}$$

where Φ is the result of mapping the *T* function in the input space depends on the dimension of the weighting vector, denoted as *w* and the bias or deviation, represented by *b*. The values of w and b are determined by minimizing the risk function described in the equation. This process involves finding the optimal coefficients that minimize the potential loss or error associated with the model. The coefficients w, b to reduce the risk function, like the following equation:

(2)

$$R = min\,\frac{1}{2}\|w\|^2 + c\frac{1}{l}\left(\left(\sum_{i=1}^{l} L_\varepsilon(y_i, f(x_i))\right)\right)$$

with limitations:

$$y_i - w\,\varphi(x_i) - b \le \varepsilon \quad (3)$$

$$w\,\varphi(x_i) - y_i + b \le \varepsilon \quad (4)$$

where $L\varepsilon$ is the loss function, w is the weight vector is the bias, R is the risk function, $||w||$ is the normalized w, ϵ is the epsilon, y_i is the actual target of training data and c is the error.

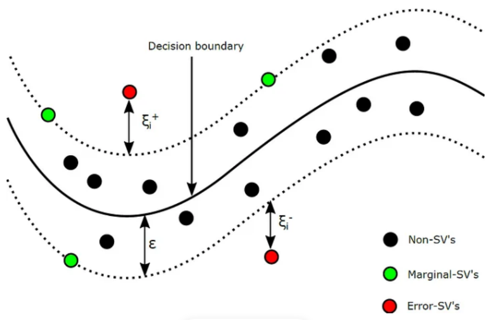

Figure 1 SVR structure (Sovia et al. 2018, IEEE Access)

3.2 *Support Vector Regression Algorithm*

The SVR approach is based on the principle of reducing error risk by estimating the function by minimising the error value. Here are the SVR technique stages for predicting the consumer price index (Erlin, 2018). Given x_i as the training data, x_j the testing data, $[R]_{ij}$ as the Hessian Matrix, $K(x_i, x_j)$ is the kernel used (Can be linear, polynomial, radial or sigmoidal), λ^2 is the scalar variable, n is the number of data, E_i is the error of i data, y_i is the actual target of training data, α_i^* is the Lagrange multipliers and α_i is the non-negative vector of Lagrange coefficient.

1. For SVR, the initialization parameters are lambda (λ), gamma (γ), complexity (C), epsilon (ε), and sigma (δ). The initialization of the initial values of the Lagrange multipliers (α_i^*) and the non-negative vector of the Lagrange coefficient (α_i) is then zero, and the maximum number of iterations is attained.

2. Using the following formula, calculate the distance between data for training data andtest data:

$$Distance = |\,x_i - x_j|^2 \quad (4)$$

3. The following equation is used to calculate the Hessian matrix:

$$[R]_{ij} = K\,(x_i, x_j) + \lambda^2 \; where\; i, j = 1, \ldots, n \quad (5)$$

4. For each set of data, i = 1… n, do the following operations:

   a. Calculating Error value

(6)

$$\boldsymbol{E}_i = \boldsymbol{y}_i - \sum_{i=1}^{n} (\boldsymbol{\alpha}^{i*} - \boldsymbol{\alpha}^{i})\,\boldsymbol{R}_{ij}$$

b. Calculate the value of $\delta\alpha^{i*}$ and $\delta\alpha^{i}$.

$$\boldsymbol{\delta\alpha^{i*}} = \boldsymbol{min}\,\{\boldsymbol{max}\,[\boldsymbol{\gamma}\,(\boldsymbol{Ei} - \boldsymbol{\varepsilon}), -\boldsymbol{\alpha}^{i*}], \boldsymbol{C} - \boldsymbol{\alpha}^{i*}\} \tag{7}$$
$$\boldsymbol{\delta\alpha^{i}} = \boldsymbol{min}\{\boldsymbol{max}[\boldsymbol{\gamma}(-\boldsymbol{Ei} - \boldsymbol{\varepsilon}), -\boldsymbol{\alpha}^{i}], \boldsymbol{C} - \boldsymbol{\alpha}^{i}\,\} \tag{8}$$

c. Calculate the value of α^{i*} and α^{i}.

$$\boldsymbol{\alpha}^{i*} = \boldsymbol{\alpha}^{i*} + \boldsymbol{\delta\alpha}^{i*} \tag{9}$$
$$\boldsymbol{\alpha}^{i} = \boldsymbol{\alpha}^{i} + \boldsymbol{\delta\alpha}^{i} \tag{10}$$

d. When convergence reaches maximum iteration or $max\,|\delta\alpha^{i*}\,| < \varepsilon$ and $max\,|\delta\alpha^{i}\,| < \varepsilon$ then the process stops. If these requirements are not satisfied, then repeat the process in step four.

5. Form the forecasting function with the following equation:

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{i=1}^{l} (\boldsymbol{\alpha}^{i*} - \boldsymbol{\alpha}^{i}\,)\,(\boldsymbol{K}\,(\boldsymbol{x}_i, \boldsymbol{x}_j) + \boldsymbol{\lambda}^2\,) \tag{11}$$

6. Once the forecasting function has been established, it is necessary to consider denormalization if the output remains in a normalized format. This involves converting the output back to its original value, ensuring it is no longer scaled or adjusted.

7. Calculate the error value using MAPE, RMSE and MAE calculation.

3.3 *Function Kernel*

A kernel is a method of transforming this two-dimensional plane into a higher dimensional space, causing it to be curved in the higher dimensional space. The kernels of SVR use various regularisation parameters to regulate the complexity and generalizability of the model. The parameters for regularisation consist of the cost parameter, degree, gamma, and epsilon.

The cost parameter, typically denoted by the letter *C*, determines the trade-off between minimising training errors and permitting margin violations. A greater value of *C* results in a narrower margin, which may result in data overfitting, whereas a lower value encourages a wider margin, forsaking training accuracy for improved generalisation. Specific to polynomial kernels, the degree parameter determines the degree of the polynomial function used to characterise the data. It regulates the complexity of the decision boundary and can be adjusted to account for varying degrees of data non-linearity. A greater degree produces a more complex decision boundary, which may result in overfitting.

Gamma, represented by the symbol γ, is a parameter used in both polynomial and radial basis function (RBF) kernels. It specifies the impact of individual training samples on the decision threshold. A higher gamma value results in a more localised and intricate decision boundary, which may overfit the training data, whereas a lower gamma value facilitates a smoother decision boundary, which improves generalisation. Epsilon, often represented as ε, determines the SVR model's error margin of tolerance. It specifies the error range within which errors are deemed permissible and do not contribute to the loss function's penalty term. A smaller epsilon results in a smaller margin, which may contribute to overfitting, whereas a larger epsilon permits larger deviations from the target and promotes a broader margin.

According to the kernel function is a function *k* where all the input vectors *x*, *z* will fulfil the following conditions:

$$\boldsymbol{k(x_i, x_j) = \emptyset(x_i)^T\, \emptyset(x)} \tag{12}$$

The four kernel functions used in the consumer price index forecasting test are as follows:

Table 1 Four types of kernels of SVR, (Ruping, 2001)

No	Kernel Function	Formulas
1	Linear	$\boldsymbol{K(x_i, x_j) = x_i^T x}$
2	Polynomial	$\boldsymbol{K(x_i, x_j) = (x_i^T x + 1)^d \; d = 1, 2, \ldots}$
3	Gaussian-Radial Basis Function	$\boldsymbol{K(x_i, x_j) = e^{(-\gamma \lVert x - x_i \rVert)^2}}$
4	Sigmoid	$\boldsymbol{K(x_i, x_j) = tanh\,(\gamma x_i^T\, x_j\; + c)}$

where $K(x_i, x_j)$ is the kernel used (Can be linear, polynomial, radial or sigmoidal), *d* is the degree, y_i is the actual value of *i* data, x_i is the training data, x_j is the testing data, x^T as the transpose of training and testing data and γ is the gamma value.

3.4 *Mean Absolute Percent Error (MAPE)*

Mean Absolute Percent Error (MAPE) is a measure of relative error. MAPE is more exact since it displays the percentage error between estimates or forecasts and actual results over a certain time period. MAPE is multiplied by 100 percent over a certain time period, and it examines the degree of variation in estimating results. This strategy is useful when the size of the forecast variable is critical for determining prediction accuracy.

The formula is:

$$\boldsymbol{MAPE = \frac{\sum_{i=1}^{n} \frac{|x_i - F_i|}{x}}{n} \times 100\%} \tag{13}$$

where x_i is the actual data, f_i is the forecast and *n* are the sample size. If the MAPE value remains below 10%, it can be concluded that the MAPE value is remarkably favourable.

3.5 *Root Mean Square Error (RMSE)*

The Root Mean Square Error represents the standard deviation of the residuals. It serves as a measure of the dispersion or variability of the residuals around the regression line. Residuals indicate how distant data points are from the regression line; RMSE indicates how widely dispersed these residuals are. In other words, it represents how closely the data clusters around the line of best fit. The root mean square error is often used to validate experimental results in climatology, forecasting, and regression analysis.

The formula is:

$$\boldsymbol{RMSE = \sqrt{\overline{(f - x)^2}}} \tag{14}$$

where x_i is the actual data, f_i is the forecast. The mean is represented by the bar above the squared differences. The same formula can be written in the slightly different notation shown below (Barnston, 1992):

$$RMSE_{f_o} = \left[\sum_{i=1}^{N} \frac{(x_i - f_i)^2}{n}\right]^{\frac{1}{2}} \tag{15}$$

where x_i is the actual data, f_i is the forecast and n is the sample size.

3.6 *Mean Absolute Error (MAE)*

Mean Absolute Error is a statistic that evaluates how close predictions or forecasts are to actual events. The accuracy measure is sensitive to scale, making it unsuitable for comparing series with different scales. The mean absolute deviation is sometimes mistaken with the mean absolute error, which is a standard metric of prediction error in time series analysis.

The formula is:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |x_i - f| \tag{16}$$

where x_i is the actual data, f_i is the forecast and n is the sample size.

4. Results and discussion

In this study, four different kernel functions of Support Vector Regression are investigated which are linear, radial, polynomial, and sigmoid. These four kernel functions are tested to determine which kernel will be better suited for the datasets and the best regularization parameter *C*. Based on the previous study and additional research on forecasting the Consumer Price Index (CPI) using Support Vector Regression (SVR), the regularization parameters used are as follows:

Table 2 The Regularization Parameters used for Different SVR Kernels

Kernel	Regularization Parameters			
	Cost Parameter, *C*	Degree, *d*	Gamma, *γ*	Epsilon, *ε*
Linear	10 100 1000	-	-	0.1
Radial		-	0.1	
Polynomial		2	-	
Sigmoid		-	0.1	

The selected regularization parameter values, such as the cost parameter (*C*) values of 10, 100, and 1000, a gamma value of 0.1, a degree value of 2, and an epsilon value of 0.1, were determined through a combination of referencing previous research papers and conducting experimentation with the specific dataset. By considering previous studies and empirical evidence that demonstrated the effectiveness of these parameter values in Consumer Price Index (CPI) forecasting using Support Vector Regression (SVR), a solid starting point was established. However, to ensure the accuracy of the parameters, an iterative process of ten trials and error was undertaken, and the pattern of plot is observed in the given dataset.

Table 3 Average of Error Measures for Training and Testing Data using SVR Kernels

Kernels	**Cost Parameter, *C***	**Training**			**Testing**		
		MAPE	RMSE	MAE	MAPE	RMSE	MAE
Linear	10	0.24	23.52	25.68	0.85	39.49	32.98
	100	0.34	24.95	25.64	0.64	37.51	31.84
	1000	**0.19**	**22.02**	**24.53**	**0.24**	**37.54**	**31.18**
Radial	10	0.23	43.40	35.63	0.20	45.58	31.67
	100	0.22	32.35	29.32	0.18	33.28	25.43
	1000	**0.20**	**25.18**	**25.30**	**0.12**	**14.49**	**14.44**
Polynomial	10	0.26	45.19	36.30	0.26	45.86	32.51
	100	0.21	25.12	28.62	**0.11**	**14.35**	**14.26**
	1000	**0.21**	**25.12**	**24.84**	0.16	15.09	15.04
Sigmoid	10	0.16	42.65	27.07	0.29	35.28	24.88
	100	0.16	35.15	25.06	0.39	38.10	32.21
	1000	**0.14**	**14.63**	**15.28**	**0.11**	**13.39**	**12.98**

Based on Table 3, the red-font digit represents the lowest error identified between different cost parameters in each SVR kernels. Comparing the kernels, the lowest error can be observed when the cost parameter, *C*= 1000. Thu, the comparison between the SVR kernels and cost parameter, *C*=1000, which provides the least error among all the four kernels, this study reveals that the SVR with a sigmoid kernel and a value of *C*=1000 gives the most accurate forecasts for the dataset of CPI Malaysia from January 2005 to December 2021. For training dataset using Sigmoid kernel, the cost parameter *C*=1000 has the lowest MAPE, RMSE, and MAE values which are 0.1435, 14.6314, and 15.2797, respectively. Similarly, for testing dataset, *C*=1000 using Sigmoid kernel has the lowest MAPE, RMSE, and MAE values which are 0.2263, 25.0902, and 24.7058, respectively. The cost parameter *C*=100 has the largest MAPE, RMSE, and MAE values for forecasting, while *C*=10 has the largest values for training dataset.

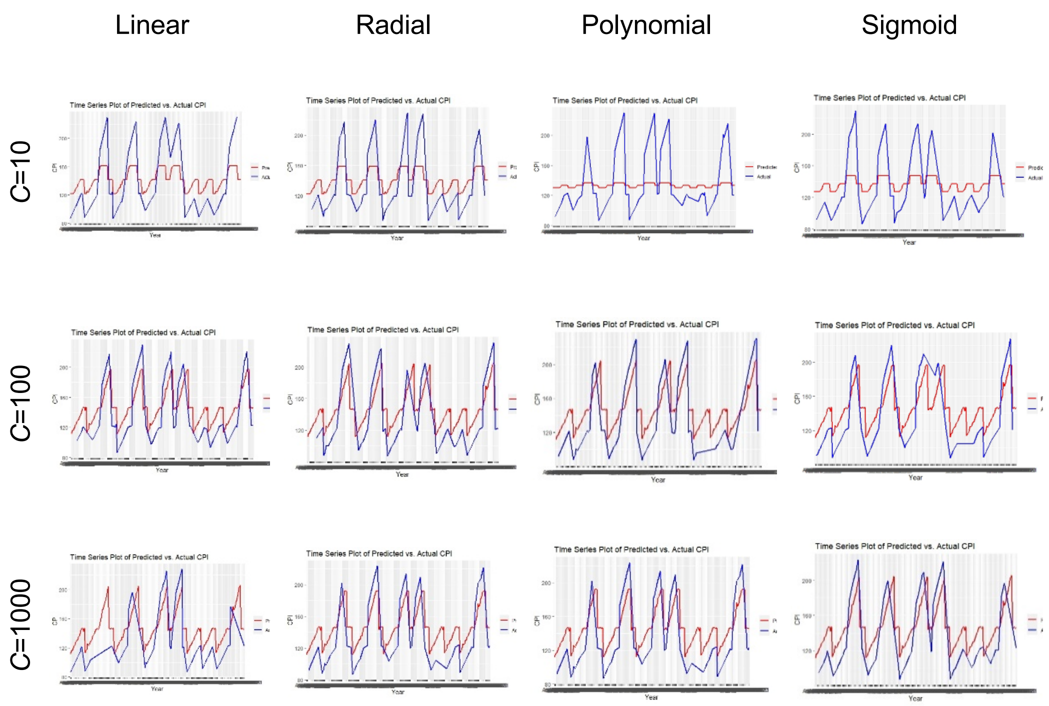

Figure 2 Comparison of Time Series Plot of SVR Kernels with Different Cost Parameters

According to Figure 2, *C* is set to 10, the Sigmoid SVR model is likely to underfit the data, leading to high bias and high training error. On the other hand, setting *C* to 100 reduces the risk of overfitting but increasing the likelihood of underfitting and slightly higher accuracy compared to *C*=10. For the Sigmoid kernel, the best value for *C* was found to be *C*=1000, with a corresponding value for gamma that maximizes the model's performance which can be seen clearly on Figure 2. This value strikes a good balance between underfitting and overfitting, resulting in low training error and good generalization performance.

Conclusion

In this study, the SVR technique was utilized with a variety of kernels and *C* values. The findings of this study indicate that the utilization of the highest cost parameter, *C* specifically*, C*=1000 value in the SVR approach yields the least error predictions for the given dataset. Comparing the kernels with *C*=1000, sigmoid kernel has the least error among all the other kernels.

The study utilized a method of averaging the outcomes of ten trials to obtain a comprehensive summary that reflects the overall efficacy and precision of the support vector regression (SVR) models in observation. The slight variations in errors indicate that the selected kernels, namely linear, radial, polynomial, and sigmoid, demonstrated proficiency in precisely approximating the fundamental data structure. The acquisition of important information and correlations enabled them to make precise predictions regarding Consumer Price Index (CPI) levels. Moreover, the dataset employed in this study, comprising monthly Malaysian CPI data, has exhibited remarkably consistent and predictable trends, thereby facilitating the generation of comparable prognostications by various kernels. If the dataset had included more intricate or nonlinear patterns, the performance discrepancies among the kernels might have been more pronounced.

Moreover, the findings of this study indicate that Support Vector Regression (SVR) is a highly effective method for forecasting Consumer Price Index (CPI) time series data. This approach holds promise as a valuable resource for scholars and economic experts seeking to generate informed projections regarding forthcoming trends. The Support Vector Regression (SVR) model demonstrated a high level of efficacy in forecasting Consumer Price Index (CPI) time series data. This can be attributed to its ability to discern intricate correlations and patterns, while not being reliant on stringent assumptions regarding the distribution of data.

References

[1] Al Ghofur, A. F., Dewi, Y. S., & Anggraeni, D. (2022). Comparison Of Support Vector Regression and Autoregressive Integrated Moving Average With Exogenous Variable On Indonesia Consumer Price Index. *SAR Journal - Science and Research*, 144-148. https://doi.org/10.18421/sar53-05

[2] Alwee, R., Shamsuddin, S. M. H., & Sallehuddin, R. (2022). Economic Indicators Selection for Property Crime Rates using Grey Relational Analysis and Support Vector Regression. *International Journal of Systems Applications, Engineering & Development, 16*, 1-8. https://doi.org/10.46300/91015.2022.16.1

[3] Barkan, O., Benchimol, J., Caspi, I., Cohen, E., Hammer, A., & Koenigstein, N. (2022). Forecasting CPI inflation components with Hierarchical Recurrent Neural Networks. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2022.04.009

[4] Cannon, A. J. (2011). Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Computers & Geosciences, 37*(9), 1277-1284. https://doi.org/10.1016/j.cageo.2010.07.005

[5] Chen, W., & Arif, M. (2022). Estimation of International Gold Price by Fusing Deep/Shallow Machine Learning. *Journal of Advanced Transportation, 2022*, 1-8. https://doi.org/10.1155/2022/6211861

[6] Dongdong, W. (2010). *The Consumer Price Index Forecast Based on ARIMA Model* 2010 WASE International Conference on Information Engineering

[7] Editya, A. S., Kurniati, N., Septianto, T., Lisdiyanto, A., & Al Haromainy, M. M. (2021). Effect of Kernel in Support Vector Regression Method to Predict Surabaya Consumer Price Index Trend 2021 IEEE 5th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)

[8] Erlin Mega Priliani, ILKOM UNNES and Anggyi Trisnawan Putra, ILKOM UNNES & Much Aziz Muslim, ILKOM UNNES (2018) Forecasting Inflation Rate Using Support Vector Regression (SVR) Based Weight Attribute Particle Swarm Optimization (WAPSO). Scientific Journal of Informatics, 5 (2). pp. 118-127. ISSN 2407-7658

[9] Glaser JI, Benjamin AS, Farhoodi R & Kording KP. The roles of supervised machine learning in systems neuroscience. Prog Neurobiol. 2019 Apr;175:126-137. doi:10.1016/j.pneurobio.2019.01.008. Epub 2019 Feb 7. PMID: 30738835; PMCID: PMC8454059.

[10] Huang, Y., Deng, C., Zhang, X., & Bao, Y. (2020). Forecasting of stock price index using support vector regression with multivariate empirical mode decomposition. *Journal of Systems and Information Technology, 24*(2), 75-95. https://doi.org/10.1108/jsit-12-2019-0262

[11] Janda, K., Kral, P., & Mikova, E. (2019). A comparison of support vector regression kernels in forecasting stock prices. Journal of Financial Management and Analysis, 32(2), 1-12.

[12] K.. -R. Muller, S. Mika, G. Ratsch, K. Tsuda & B. Scholkopf, "An introduction to kernel-based learning algorithms," in IEEE Transactions on Neural Networks, vol. 12, no. 2, pp. 181-201, March 2001, doi: 10.1109/72.914517.

[13] Khamis, A. B. (2020). Time Series Forecasting of Malaysia Producer Price Index using ARIMA and Grey Models. *International Journal for Research in Applied Science and Engineering Technology, 8*(8), 1343-1348. https://doi.org/10.22214/ijraset.2020.31160

[14] Momo, S. I., Riajuliislam, M., & Hafiz, R. (2021). Forecasting of Inflation Rate Contingent on Consumer Price Index: Machine Learning Approach. In *Intelligent Computing and Innovation on Data Science* (pp. 137-144). https://doi.org/10.1007/978-981-16-3153-5_17

[15] Mwanga, Y. (2020). Arima Forecasting Model for Uganda's Consumer Price Index. *American Journal of Theoretical and Applied Statistics, 9*(5). https://doi.org/10.11648/j.ajtas.20200905.17

[16] Qin, X., Sun, M., Dong, X., & Zhang, Y. (2018). *Forecasting of China Consumer Price Index Based on EEMD and SVR Method* 2018 2nd International Conference on Data Science and Business Analytics (ICDSBA),

[17] Qiu, Y. (2020). Forecasting the Consumer Confidence Index with tree-based MIDAS regressions. *Economic Modelling, 91*, 247-256. https://doi.org/10.1016/j.econmod.2020.06.003

[18] Qu, F., Wang, Y.-T., Hou, W.-H., Zhou, X.-Y., Wang, X.-K., Li, J.-B., & Wang, J.-Q. (2022). Forecasting of Automobile Sales Based on Support Vector Regression Optimized by the Grey Wolf Optimizer Algorithm. *Mathematics, 10*(13). https://doi.org/10.3390/math10132234

[19] Riofrío, J., Chang, O., Revelo-Fuelagán, E. J., & Peluffo-Ordóñez, D. H. (2020). Forecasting the Consumer Price Index (CPI) of Ecuador: A Comparative Study of Predictive Models. *International Journal on Advanced Science, Engineering and Information Technology, 10*(3). https://doi.org/10.18517/ijaseit.10.3.10813

[20] Sovia, Rini & Yanto, Musli & Gema, Rima & Fernando, Rizki. (2018). Bank Indonesia Interest Rate Prediction and Forecast With Backpropagation Neural Network. 429-435. 10.1109/ICITSI.2018.8695914.

[21] Venkadasalam, S. (2015). The Determinant of Consumer Price Index in Malaysia. *Journal of Economics, Business and Management, 3*(12). https://doi.org/10.7763/joebm.2015.V3.344

[22] Wang, W., Zhang, Z., Wang, L., Zhang, X., & Zhang, Z. (2022). Mixed-frequency data-driven forecasting the important economies' performance in a smart city: a novel RUMIDAS-SVR model. *Industrial Management & Data Systems, 122*(10), 2175-2198. https://doi.org/10.1108/imds-01-2022-0014

[23] Wong, J. M. W., & Ng, S. T. (2010). Forecasting construction tender price index in Hong Kong using vector error correction model. *Construction Management and Economics, 28*(12), 1255-1268. https://doi.org/10.1080/01446193.2010.487536

[24] Xiao, J., Wang, M., Tian, L., & Zhen, Z. (2018). The measurement of China's consumer market development based on CPI data. *Physica A: Statistical Mechanics and its Applications, 490*, 664-680. https://doi.org/10.1016/j.physa.2017.08.135

[25] Zahara, S., & Sugianto. (2020). *Multivariate Time Series Forecasting Based Cloud Computing For Consumer Price Index Using Deep Learning Algorithms* 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI),

[26] Zhang, L., Wang, Y., Zhang, Y., & Hu, Y. (2019). A comparison of support vector regression kernels in forecasting the inflation rate of China. Journal of Applied Statistics, 46(9), 1667-1686.

[27] Zhang, Y., Li, W., Li, W., & Wang, S. (2019). Comparison of support vector regression kernels in forecasting the inflation rate of China. PLOS ONE, 14(11), e022