# Improving Cryptocurrency Price Prediction Accuracy with Multi-Kernel Support Vector Regression Approach

**Subba Reddy Thumu [a, *], Geethanjali Nellore [a]**

[a] Department of Computer Science & Technology, Sri Krishnadevaraya University, Anantapur, India
*Corresponding Author Email: subbareddy.thumu@gmail.com

**Abstract:** Cryptocurrencies are digital assets that have attracted a lot of investment and attention. It is challenging and essential for investors and traders to predict their stock price movements. Making accurate predictions about cryptocurrency prices is crucial for avoiding losses and gaining profits. Our research proposes a novel method for predicting the stock closed prices of three popular cryptocurrencies: Bitcoin, Ethereum and Polkadot. The SVR (Support vector regression) machine learning method can provide robust and accurate predictions for nonlinear and nonstationary data. This paper compares SVR radial basis functions (RBFs) and hybrid kernels based on cryptocurrency data characteristics. SVR parameters such as regularization, gamma, and epsilon can also be tuned using grid search. Our approach is tested on real-world cryptocurrency stock prices collected from Yahoo Finance. Prediction performance is measured using regression metrics like MAPE (Mean absolute percentage error) and $R^2$ score. In our work, a MAPE value of 0.07772 and an $R^2$ score of 0.9999 have been obtained. The results of our experiments indicate that our approach is significantly more accurate and reliable than existing methods.

**Keywords:** Cryptocurrency, Stock Prediction, SVR, RBF, Hybrid Kernels, Grid Search, Regression Metrics

## 1. Introduction

Cryptocurrency markets are notoriously volatile and complex, so robust predictive models are in high demand to help investors and researchers navigate this dynamic environment [1, 2]. Online transactions are anonymous and secure with cryptocurrencies. Since cryptocurrencies are decentralized, central control over them has been greatly reduced, resulting in an impact on international trade and relations [3]. An innovative approach is used in this work to refine cryptocurrency price predictions. As more investors seek lucrative opportunities in the cryptocurrency market, Bitcoin (BTC), Ethereum (ETH), and Polkadot (DOT) have become focal points in global financial markets.

BTC transactions are verified and recorded in a public ledger (the blockchain), eliminating the need for central intermediaries or trusted record-keeping authorities [4]. Transaction blocks contain SHA-256 cryptographic hashes of previous transaction blocks, making them immutable records of all transactions [5]. The blockchain-based ETH platform enables developers to create decentralized applications (DAPs) and is used as the ETH network's currency for transaction fees. Cryptocurrency payment platform ETH allows users to complete transactions by sending requests to machines that complete the requested actions [6]. The DOT network features parachains and user-created blockchains that can be customized while benefiting from the same security measures. Additionally, parachains offload much of Polkadot's processing load. DOT is considered a potentially competitive and innovative project [7]. Nominators financially support validators as a sign of trust in their integrity.

In December 2023, the global cryptocurrency market cap was $1.74 Trillion, a 109.3% change from a year ago [8]. With a market cap of $836 billion, BTC dominates 48.03%. Based on 11887 cryptocurrencies tracked across 959 exchanges, the chart below shows the total market cap and volume of cryptocurrencies. Based on the top 10 ranking, the graph shows BTC's dominance percentage.

In order to make strategic decisions, it is essential to be able to forecast their price movements accurately. Cryptocurrency markets often feature intricate patterns and sudden shifts that are difficult to capture with traditional time series analysis tools [9, 10]. SVR (Support vector regression) has shown promising results in this context. In this work, multi-kernel strategies and grid search optimization (GSO) are combined to enhance SVR prediction accuracy [11, 12]. Using radial basis function (RBF) kernels for SVR with cryptocurrency price dynamics has been widely adopted.

**Figure 1.** Cryptocurrency stock shares percentage

(Source: https://www.coingecko.com/en/global-charts)

However, new kernels are incorporated in this work, tailored to address nuances in cryptocurrency price dynamics (as shown in Figure 1). Incorporating diverse kernels improves the model's ability to adapt to market fluctuations. In addition, GSO seeks out the most effective configurations for the SVR model by systematically exploring hyperparameter combinations. Predictive models will be evaluated using comprehensive regression metrics, including $R^2$ scores and mean absolute percentage errors (MAPE) [13, 14]. These metrics are used as benchmarks to assess how well a model captures and predicts cryptocurrency stock closing prices.

The present work aims to provide useful insights into cryptocurrency price prediction, the literature, results, and methodology. SVR models will be fine-tuned with multi-kernel strategies and parameter optimization to provide cryptocurrency practitioners with better tools to navigate the intricate landscape of cryptocurrency investments [15].

The rest of the paper is structured as follows. In Section 2, the related work is summarized; in Section 3, the concepts involved in this research are explained; Section 4 proposes an SVR approach with multiple kernels; An analysis of the experimental results is also presented and the final section presents conclusions and recommendations for future research.

## 2. Related Work

This section discusses various machine learning models based on SVR and kernels applied to cryptocurrency datasets that will be used to construct the proposed model.

The study by Kate Murray *et al.* [1] uses statistical, machine learning, and deep learning approaches to predict five popular cryptocurrencies' prices: XRP, BTC, litecoin, ETH, and monero. In this paper, the author uses a unified framework to run extensive experiments on a dataset containing daily prices and volumes for five cryptocurrencies from January 2016 to December 2019. RMSE, MAPE, and $R^2$ score accuracy are used to assess prediction performance in the paper. This study found that the long-short-term memory (LSTM) model outperformed statistical and machine learning models, including the SVR model, most of the time.

According to Helder Sebastiao and Pedro Godinho [2], cryptocurrencies BTC, ETH, and Litecoin exhibit a high degree of predictability. The study also investigates machine learning-based trading strategies, including linear models, Random Forest, and SVR with RBF kernels. GSO is also used to tune the SVR model's hyperparameters. The SVR model outperforms existing methods using the new kernels and achieves high prediction accuracy.

In Fan Fang *et al*. [9], a comprehensive analysis of cryptocurrency trading research is presented, including analysis of datasets, research trends, research objects, and technologies, concluding with some promising opportunities to trade cryptocurrency currently available. In addition to reviewing 146 papers on cryptocurrency trading, the author examines several that use SVR and machine-learning techniques to predict cryptocurrency prices.

According to Deny Haryadi *et al*. [16], their SVR model outperforms the risk of investing in cryptocurrencies. DOT cryptocurrency stock price is

predicted using linear kernels and RBF kernels with min-max normalization. After hyperparameter tuning with GSO, the RBF kernel gives better results based on regression metrics model accuracy and MAPE.

In a bivariate time-series method using the Morgan Stanley Capital International (MSCI) World Index as a predictor variable, Saad Ali Alahmari [17] applies SVR with linear, polynomial, and RBF kernels to predict the prices of BTC, XRP, and ETH. Based on three regression metrics: coefficient of determination (R² score), root mean squared error (RMSE), and MAPE, the author evaluates the SVR model's performance: A significant predictor of cryptocurrency prices is the MSCI World Index, based on the SVR model with the RBF kernel.

Barcenas *et al.* (2022) [18] use of voice signal analysis as an objective measure of PD progression, the Unified Parkinson's Disease Rating Scale (UPDRS), has its limitations. The researchers proposed a Support Vector Regression (SVR) model with a combined kernel function. This model captures both global and local information in the data, which is a crucial requirement for SVR models. The best performance was achieved when the kernel function combined a radial basis function and a polynomial basis function. It revealed non-linear relationships between voice features and UPDRS scores, providing deeper insights into the disease progression, also it achieved significantly improved prediction performance compared to existing methods. By incorporating factors like age and gender, the model

could potentially describe the dynamics of UPDRS changes based on patient monitoring data.

This research seeks to contribute to the existing knowledge base by synthesizing insights from these diverse areas of literature and proposing a comprehensive framework for fine-tuning cryptocurrency predictions, which integrates SVR with multi-kernel strategies and GSO. Combining these elements will produce a more nuanced and effective method for forecasting cryptocurrency prices.

## 3. Data & Methodology

### 3.1 Dataset

The proposed work involved gathering historical daily cryptocurrency stock prices for Bitcoin(BTC), Ethereum(ETH), and Polkadot(DOT) [16, 17, 19], comprising date, open price, high price, low price, close price, adjusted close price, and volume. The datasets used for the work were gathered from Yahoo Finance. Microsoft Windows 11 x64 was installed on a computer with 16 GB of RAM and an Intel core i7 processor. Implementations of each algorithm were carried out using the MATLAB programming language.

Table 1 shows the essential information for these datasets. Figure 2 shows the chronological dataset for three cryptocurrencies (BTC, ETH and DOT) for the selected time series. Based on the selected cryptocurrencies, the models were trained and extracted the following data:
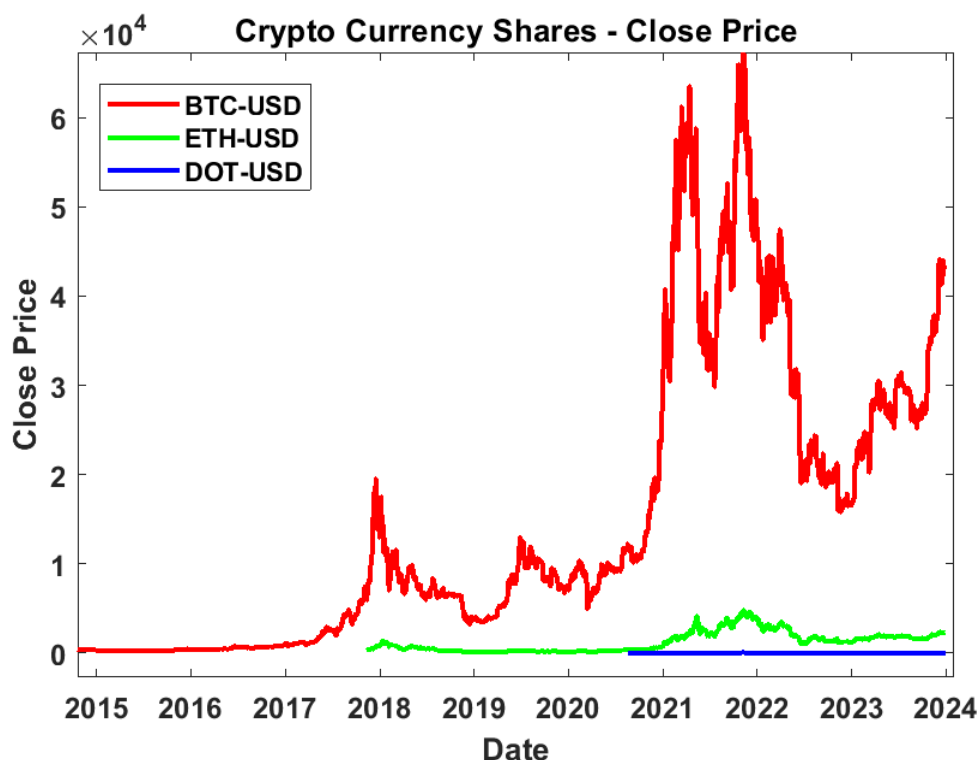


**Figure 2.** Cryptocurrency close prices of BTC, ETH, and DOT

**Table 1.** Basic information about the Datasets

| Name | Release Year | Current Price | Market Cap | 24h Volume |
|---|---|---|---|---|
| Bitcoin (BTC) | 2009 | $45,292.37 | $887,038,348,447 | $31,761,236,402 |
| Ethereum (ETH) | 2015 | $2,375.11 | $285,454,275,223 | $11,610,569,476 |
| Polkadot (DOT) | 2020 | $8.54 | $10,795,645,855 | $314,075,450 |

1) Cryptocurrency of Bitcoin (17-09-2014 to 30-12-2023).

2) Cryptocurrency of Ethereum (10-11-2017 to 30-12-2023).

3) Cryptocurrency of Polkadot (20-08-2020 to 30-12-2023).

## 3.2 Feature Extraction

Among the collected datasets, BTC, ETH, and DOT contain 3392, 2242, and 1228 records, respectively. The datasets were applied on multi-kernel strategies and grid search optimization (GSO) are combined with SVR Machine-learning models (SVR - CPR, SVR - Hybrid APR). As a result of the features presented, the closing price of a stock is a pivotal attribute in stock price predictions [20, 21]. It is a target attribute for machine learning models, encapsulating investors' sentiment. There were two sets of data: a train set containing 80% and 70% of the data, while a test set containing 20% and 30%. The training model has not seen the test dataset, which aligns perfectly with the definition of forward testing. A model's strong performance on a significant amount of unseen data suggests a lower risk of overfitting.

## 3.3 Support Vector Regression

An algorithm called the SVR algorithm [22, 23] uses machine learning to predict continuous output values based on input values. Essentially, it aims to minimize prediction error by placing most data points within the hyperplane, handling outliers, and overfitting and underfitting. SVR's optimization function can be defined as follows:

$$\min_{w,b,\xi_i,\xi_i^*} \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \right\} \qquad (1)$$

$w$ and $b$ are the undetermined constant vectors, where $C$ is a penalty coefficient parameter controlling the severity of penalized loss when a training mistake occurs. Slack variables $\xi_i$ and $\xi_i^*$ were included to account for training errors above and below an interval of $\varepsilon$ tube.

Several kernel functions are widely used in SVR, including linear, polynomial, RBF, and sigmoid kernels [20]. As a result, the decision function for SVR may be written as follows:

$$f(x) = \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(x_i, x') + b \qquad (2)$$

where $\alpha_i$ and $\alpha_i^*$ are Lagrange multipliers with kernel function K, where $K(x_i, x') = \Phi(x_i) \cdot \Phi(x')$.

A kernel function is used to map input data to a higher-dimensional space, making SVR capable of handling nonlinear relationships between input variables and targets. It can capture more complex patterns in data by determining the similarity between input vectors. Different kinds of kernel functions exist with different formulas. Kernel functions include the following [14, 24, 25].

Linear kernel of degree *n* is:

$$K_{Linear}(x, x') = x^T x' \qquad (3)$$

Polynomial kernel of degree *d* is:

$$K_{Poly}(x, x') = (\gamma \cdot x^T x' + c)^d, \gamma > 0 \qquad (4)$$

Radial basis function kernel is:

$$K_{RBF}(x, x') = \exp(-\gamma . \|x - x'\|^2) \qquad (5)$$

Sigmoid kernel is:

$$K_{Sig}(x, x') = tanh(\gamma . x^T x' + c) \qquad (6)$$

ANOVA kernel is:

$$K_{ANOVA}(x, x') = \sum_{i=1}^{n} \exp(-\gamma . (x_i - x_i')^2)^d \qquad (7)$$

Here $c, d$ and $\gamma$ are the kernel parameters, and $x$ and $x'$ are support vectors and test vectors of dataset [26].

A variety of kernels are applied with relevant parameters in this research. Parameter values must be experimented with to find the optimal ones. A different default value is used when calling the SVR function in RBF, sigmoid, and polynomial kernels; if these parameters ('γ' and 'C') are not set optimally, default values will be used. The regularization parameter 'C' indicates an empirical error penalty in the dataset. A 'γ' parameter describes how the kernel spreads along the data points for nonlinear hyperplanes. A lower 'γ' value and a higher C value indicate a larger decision range. The support vector regression model is represented in figure 3 for better understanding.

## 3.4 Grid Search Optimization

The selection of hyperparameters for SVR models has a significant impact on the identification of results [27].
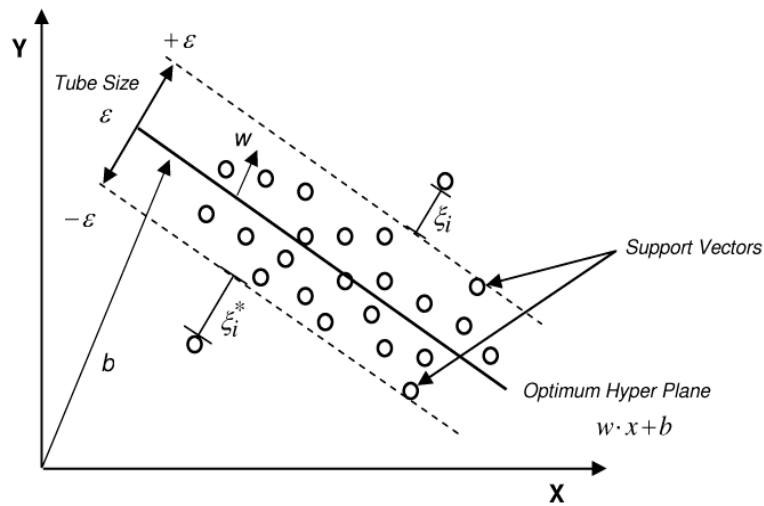
**Figure 3.** Support Vector Regression model

As the parameters are nonlinear, it is often necessary to run numerous experiments to determine the combination of parameters, such as penalty parameter 'C' and kernel parameter 'γ' [28]. When hyperparameters are searched using a grid, an exhaustive search is performed using a subset of parameters.

A hyper parameter is defined by its minimum (lower bound), maximum (upper bound), and number of steps. Using grid search, the optimal parameters can be determined by dividing the range of parameters into grids and crossing them all. Cross-validation is used as a performance measure to optimize the SVR parameter in a grid search. An objective of grid search is to find the optimal hyper parameter combination to avoid overfitting and under fitting problems and accurately predict unknown data. Below, you'll find the pseudocode for the grid search algorithm. In a grid search across different hyper parameter values, five values are considered for each hyper parameter, so 25 different combinations are evaluated and compared. For parameter 'C', 0.1, 1, 10, 100, and 1000 were considered optimal values, and for parameter gamma 'γ', 1, 0.1, 0.01, 0.001, and 0.001 values.

### 3.5 Evaluation Metrics

Regression metrics serve as quantitative measures to assess the goodness-of-fit of regression models and quantify the errors between predicted and actual values. These metrics play a crucial role in evaluating the performance of predictive models that focus on continuous numeric outcomes. Calculating the MAPE and the $R^2$ score will determine which model produces the most accurate results for the original time series [13, 14].

1. MAPE measures the average of absolute error between predicted values and actual values. It provides insights into how well the model predicts relative to the actual values. A lower MAPE indicates better prediction accuracy.

$$MAPE = \frac{1}{M}\sum \frac{|z_i - \hat{z}_i|}{z_i} \tag{8}$$

2. $R^2$ score measures how well the explanatory variables (features) explain the variance in the dependent variable (target). It assesses the goodness of fit of the regression model. $R^2$ score ranges from 0 to 1 and higher value indicates a better fit of model.

$$R^2\ score = 1 - \frac{SSR}{SSM} = 1 - \frac{\sum(z_i - \hat{z}_i)^2}{\sum(z_i - \bar{z}_i)^2} \tag{9}$$

SSR = Squared sum error for Regression line, SSM = Squared sum error for Mean line.

Where $M$ is the total number of records, $z_i$ − Actual value, $\hat{z}_i$− Predicted value, and $\bar{z}_i$ − Mean value of $z_i$

## 4. Proposed Method

To develop custom kernels for SVR to enhance model performance, capture complicated data patterns, incorporate domain knowledge, and enable users to develop customized solutions for specific problem domains. Custom kernels can be created by combine different kernels or by create new kernels that are not included in the standard library. Specifying a custom kernel in SVR as a function that takes two input vectors as arguments and returns a scalar value as its output is possible. The function must be positive definite to be continuous and symmetric. The kernel matrix can also be pre-computed and passed to the SVR algorithm as input data.

Standard RBF kernels in SVR excel at capturing broad trends, but struggle with complex non-linear relationships in stock prices.
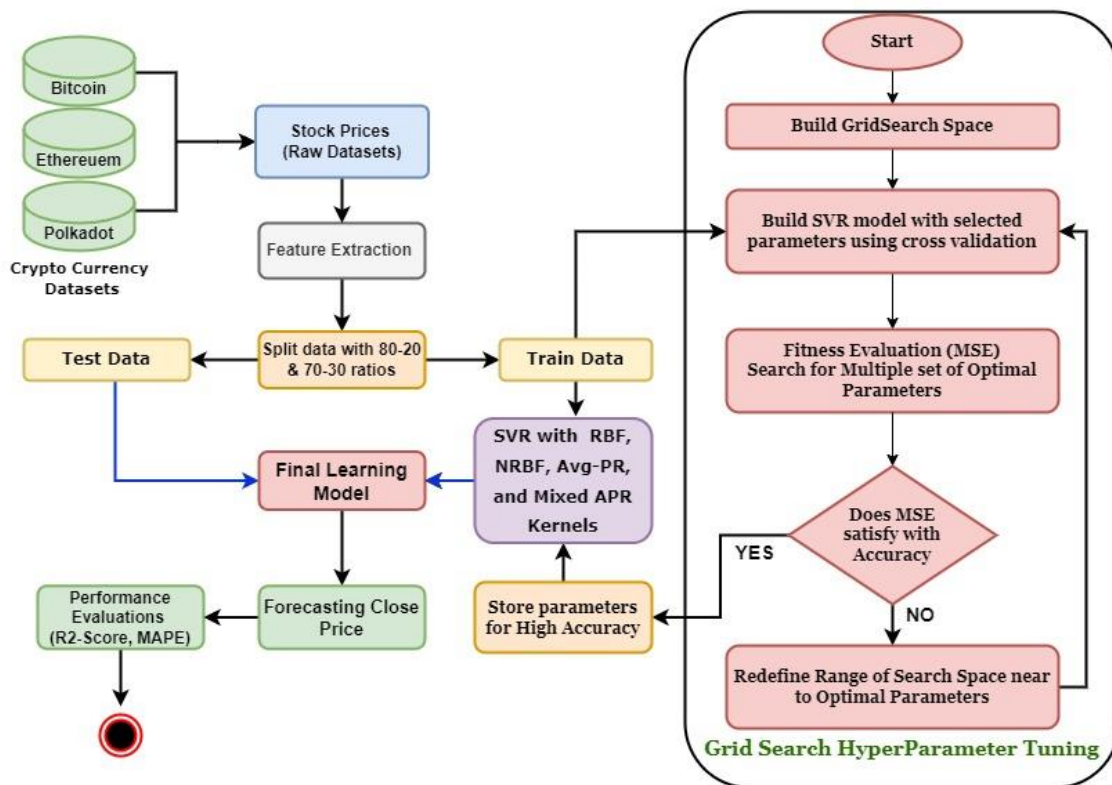
**Figure 4.** An improved SVR model for predicting cryptocurrency stocks

Polynomial kernels, while more adaptable, can be challenging to fine-tune and prone to overfitting. A Combined Polynomial RBF (CPR) kernel offers a solution. It leverages the strengths of both, capturing both global trends and localized patterns, potentially improving flexibility and boosting predictive performance. However, the increased complexity of CPR kernels necessitates careful consideration of the trade-off between adaptability and interpretability.

A combination of ANOVA, RBF, and Polynomial kernels offers several advantages over a standard RBF kernel for stock price prediction with SVR. This trio directly addresses key characteristics of stock prices. RBF captures global trends & broad trends, Polynomial handles non-linear relationships between features, and ANOVA excels at interaction effects, crucial for how various economic and company-specific events influence each other.

By combining these strengths, the hybrid APR (Anova Polynomial RBF) kernel model can potentially achieve a more comprehensive understanding of the factors driving stock prices. This can lead to improved flexibility, capturing both broad trends and intricate interactions, and potentially better prediction performance compared to a sole RBF kernel.

$$\text{CPR Kernel} = (K_{Poly} + K_{RBF}) / 2 \qquad (10)$$

$$\text{Hybrid APR Kernel} = \alpha.K_{RBF} + (1 - \alpha).(K_{Anova} + K_{Poly}) \qquad (11)$$

Where **α** is kernel parameter.

The proposed model uses SVR and GSO to forecast cryptocurrency stock close prices [29]. The mathematical formulation of SVR and optimization methods is presented in Section 3. Figure 4 illustrates the stages that make up a suggested model for predicting stock prices using datasets. These steps can be applied to the work in MATLAB.

1. Make a selection of datasets and perform feature selection

2. Prepare training and validation datasets with 80:20 and 70:30 split ratios.

3. SVR parameters should be tuned by GSO to yield high accuracy while saving time

4. Conduct training and validation of the proposed SVR model using high accuracy parameters

5. Analysing the performance of the proposed models based on its results

**Algorithm**

BEGIN

1. Read the dataset

2. Using 80:20 and 70:30 ratios, divide the dataset into train and test/validation datasets.

3. Repeat the steps below for every type of kernel.

Set the hyperparameters of the SVR model like kernel scale, C, γ, and ε

```
Set i=1, j=1

while i < imax

while j < jmax

Choose the model hyperparameters that perform best
based on MSE.

Obtain the SVR prediction function using these
hyperparameters.

j= j + 1

end while

i = i + 1

end while

4. Extract optimized hyperparameters based on MSE
and use these hyperparameters to build an SVR
model.

5. Regression analysis should be done to determine
the optimum SVR kernel model.

END
```

## 5. Results and Discussion

In the BTC and ETH datasets, 80% and 20% of the datasets have been selected for training and testing. For SVR machine learning, the author applies linear, polynomial, and RBF kernel functions to these datasets [17]. Consequently, the RBF kernel function achieved better regression results, with 78% for BTC and 56% for ETC.

In order to improve the results, the DOT dataset has been subjected to a min-max normalization. Later, training and testing datasets were split with an 80% and 20% ratio. Using a GSO technique for tuning the hyper parameters, the authors used a variety of kernel functions, including linear and RBF, for the SVR machine-learning model [16]. Thus, RBF kernels achieve a better regression result with a 5.289 MAPE and 90% $R^2$ score.

Statistical significance is crucial for evaluating model performance, the ANOVA (Analysis of Variance) test on our compared kernel methods (RBF, NRBF, CPR, and Hybrid APR) across BTC, ETH, and DOT datasets (with p-values of 0.5168, 0.2846, and 0.1262, respectively are shown in Figure 5) didn't reveal statistically significant differences in performance (measured by MAPE or $R^2$ scores). Although the p-value for the DOT dataset (0.1262) hints at a potential trend, none of the p-values fell below a commonly used threshold (e.g., 0.05) to definitively reject the null hypothesis of no difference between the methods. This suggests that further investigation might be needed to identify statistically significant performance variations between the kernel methods across these datasets.

Through our datasets, we determined which parameters ('γ' and 'C') impact the SVR model the most. We got better results with more points constrained in the proposed kernels with respective optimal gamma ('γ') values ranged from $10^{-5}$ to $10^1$. Consequently, the optimal empirical error penalty ('C') ranged from $10^{-2}$ to $10^4$ with step size of 0.1. The 'C' parameter value is high, so the model is penalized heavily for unstable data. However, a lower 'γ' value may indicate that the decision region has been enlarged. Furthermore, a parameter called epsilon controls the loss functions used in the SVR, allowing the optimal value to be found.

```
Source        SS          df        MS          F      Prob>F
---------------------------------------------------------------
Groups    8.50603e+08        4    2.12651e+08   0.81   0.5168
Error     4.43645e+12    16955     2.6166e+08
Total      4.4373e+12    16959
```

```
Source        SS          df        MS          F      Prob>F
---------------------------------------------------------------
Groups    6.27879e+06        4     1.5697e+06   1.26   0.2846
Error      1.3994e+10    11205    1.24891e+06
Total     1.40003e+10    11209
```

```
Source        SS          df        MS          F      Prob>F
---------------------------------------------------------------
Groups       1047.5          4      261.877    1.8    0.1262
Error       893532.2      6135      145.645
Total       894579.7      6139
```

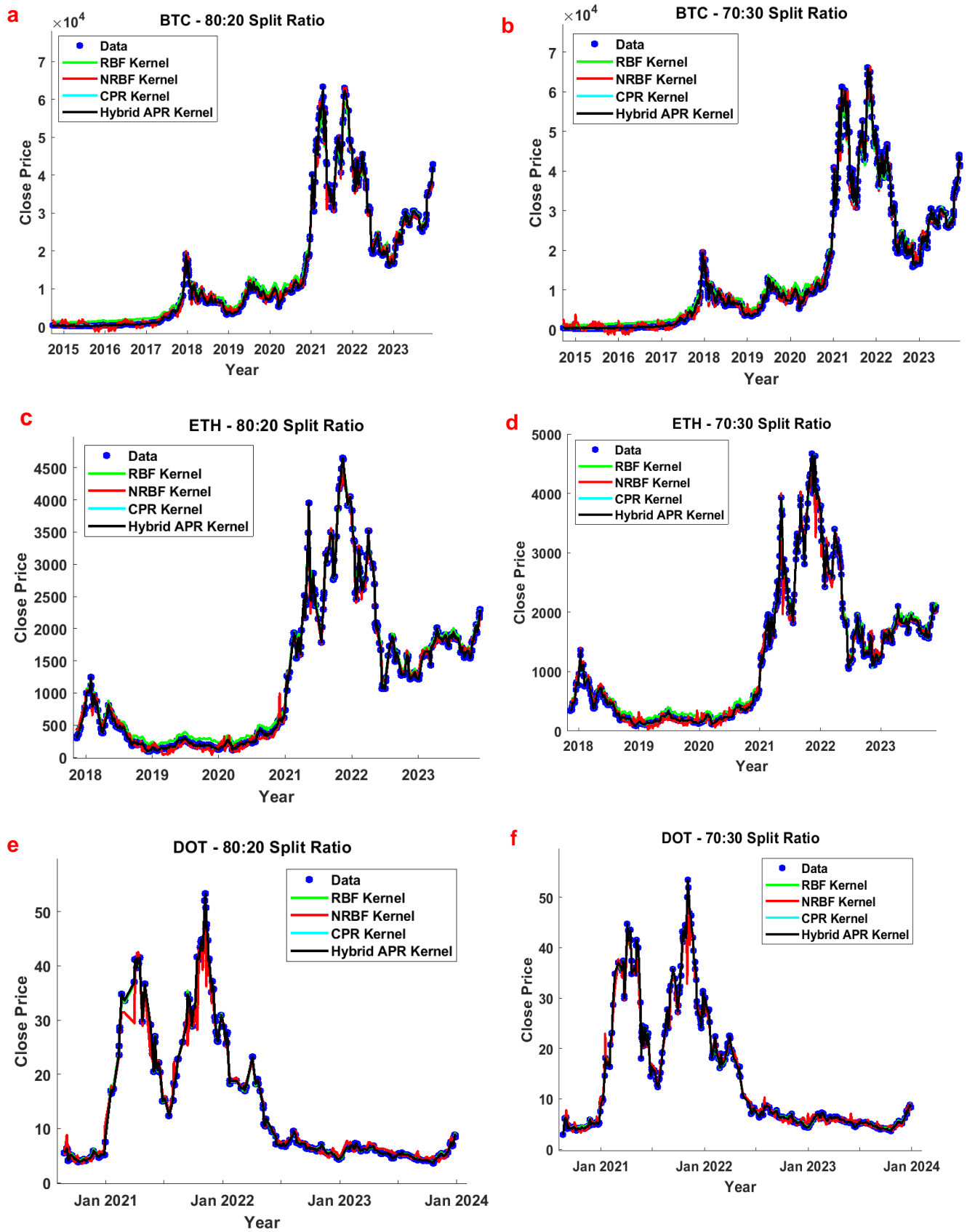**Figure 5.** ANOVA test results for the BTC, ETH and DOT datasets

**Figure 6 (a).** Actual vs Predict graph for BTC 80:20 split data, **(b),** Actual vs Predict graph for BTC 70:30 split data, **(c),** Actual vs. Predict graph for ETH 80:20 split data, **(d),** Actual vs Predict graph for ETH 80:20 split data, **(e),** Actual vs Predict graph for DOT 80:20 split data, **(f),** Actual vs Predict graph for DOT 80:20

**Table 2.** Comparison of various kernels results with 80:20 split ratio

| Kernels with Grid Search and Cross-validation | $R^2$ score (Accuracy) | | | MAPE | | |
|---|---|---|---|---|---|---|
| | BTC | ETH | DOT | BTC | ETH | DOT |
| RBF Kernel [17] | 0.95541 | 0.99521 | 0.99525 | 52.026 | 14.176 | 7.508 |
| NRBF Kernel [16] | 0.964891 | 0.998554 | 0.99833 | 29.5357 | 3.941979 | 5.98995 |
| CPR Kernel | **0.9997** | **0.9999** | **0.9999** | **1.0999** | **0.29172** | **0.07772** |
| Hybrid APR Kernel | 0.9969 | 0.9988 | 0.9998 | 1.7451 | 0.52169 | 0.27917 |
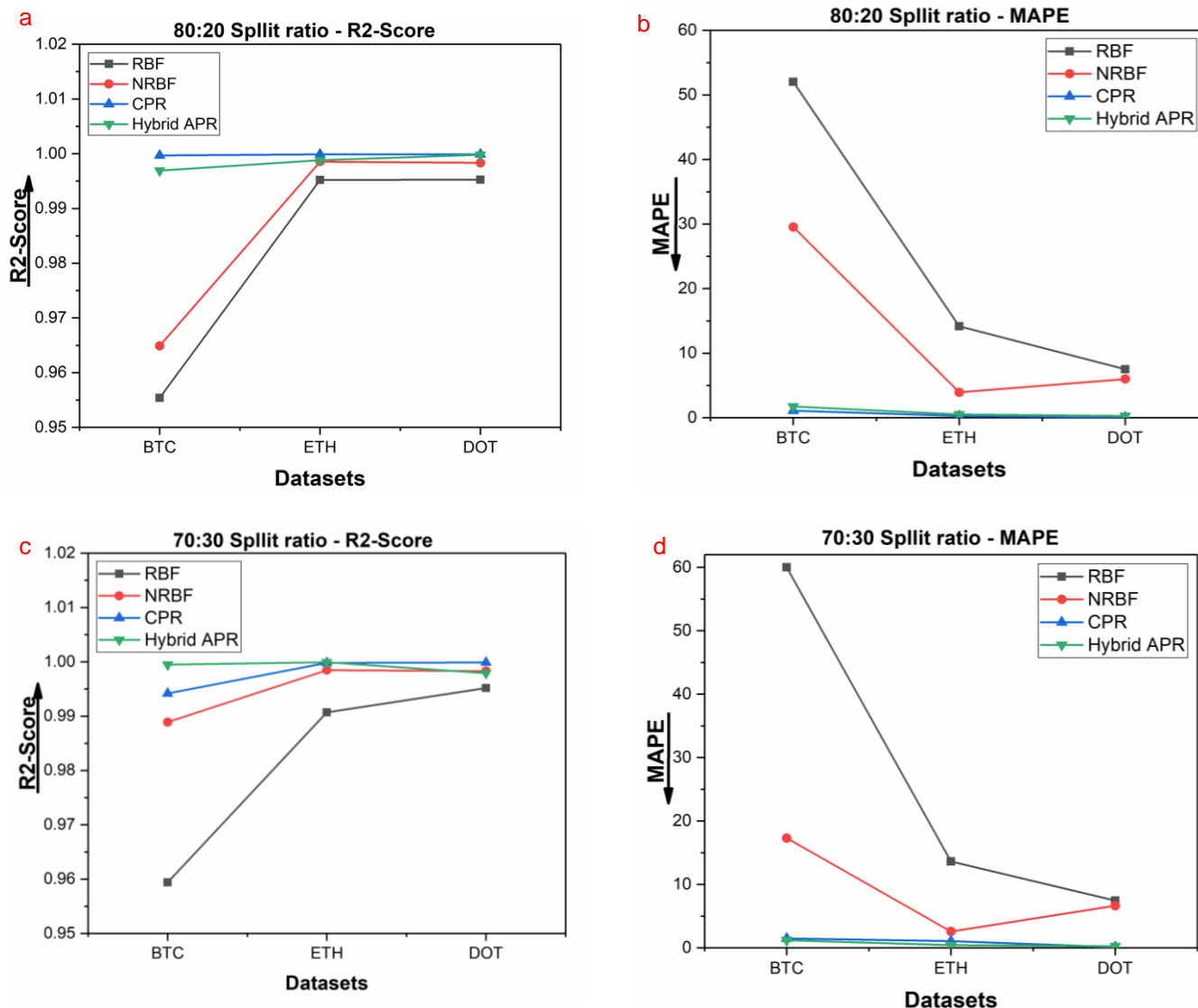


**Figure 7(a).** $R^2$-Score metrics for 80:20 split ratio, **(b),** MAPE metrics for 80:20 split ratio, **(c),** $R^2$-Score metrics for 70:30 split ratio, **(d),** MAPE metrics for 70:30 split ratio

**Table 3.** Comparison of various kernels results in a 70:30 split ratio

| Kernels with Grid Search and Cross-validation | $R^2$ score (Accuracy) | | | MAPE | | |
|---|---|---|---|---|---|---|
| | BTC | ETH | DOT | BTC | ETH | DOT |
| RBF Kernel [17] | 0.95939 | 0.99071 | 0.99519 | 60.002 | 13.609 | 7.4399 |
| NRBF Kernel [16] | 0.988905 | 0.998474 | 0.99831 | 17.30328 | 2.575257 | 6.63771 |
| CPR Kernel | 0.9942 | 0.9998 | **0.9999** | 1.4901 | 1.0604 | **0.10894** |
| Hybrid APR Kernel | **0.9995** | **0.99991** | 0.9979 | **1.2092** | **0.44444** | 0.26336 |

Datasets [30-32] for training and testing include BTC, ETH, and DOT with split ratios of 80:20 and 70:30. By using the holdout partitioning method, the dataset has been divided into train and test datasets. A random sample was then selected from the entire dataset for both the train and test datasets [33]. As part of an 80:20 split ratio, the test dataset contains 678, 448, and 245 records. The training dataset contains 2714, 1794, and 983 records. For a 70:30 split, there are 2375, 1570, and 860 records for trains and 1017, 672, and 368 for test data for BTC, ETH, and DOT stocks. The regression metrics chosen are MAPE and $R^2$ score.

SVR models use grid search to fit hyper parameters to RBF, NRBF (Normalized RBF), CPR, and Hybrid APR kernels. By using these results, regression metrics (Table 2 & Table 3) can be derived, and predicted graphs can be generated (Figures 6(a) - 6(f)) based on these results. The proposed hybrid APR and CPR kernels are compared with the RBF kernels, which are taken from the normal data set [17], and the other one is taken from the min-max normalized dataset as the training and test data set for the RBF kernels [16].

For these ratios of 80:20 and 70:30, the proposed model outcomes are represented by predicted graphs. The prediction curve graph consists of four kernel curves and one data curve. This analysis allows us to understand how model complexity affects performance. All the datasets up to the year 2021 (Figures 7(a) - 6(f)) show rapid stock growth followed by sudden drops and increases. The proposed kernels CPR and Hybrid APR can accurately predict the rapid changes for the entire random test data, while the RBF and NRBF kernels cannot do so. Models must balance under fitting and overfitting to perform well on tests and training.

## 5.1 Comparison & Discussion

Based on regression metrics MAPE and $R^2$ score (shown in Figure 6), RBF [17], NRBF [16], CPR and hybrid APR kernels have been compared across all datasets with split ratios of 80:20 and 70:30. Even though NRBF kernels use SVR with min-max normalization techniques with ranges of 0 to 1, the model does not produce accurate predictions.

CPR and Hybrid APR kernel methods are compared with RBF and NRBF kernel methods to predict the close price (Tables 2 & 3). A CPR kernel with an 80:20 split ratio achieves the lowest MAPE values for BTC, ETH, and DOT datasets at 1.0999, 0.29172, and 0.07772. Regarding accuracy comparison, the proposed kernel CPR kernel is achieving high $R^2$ score values of 0.9997, 0.9999, and 0.9999. Similarly, for 70:30 split ratios, the hybrid APR kernel achieved better values for the BTC and ETH datasets, and the CPR kernel got better results for the DOT dataset. Based on all these observations, the CPR kernel predicts appropriately

when the training dataset is huge, and when the dataset size is less, the hybrid APR kernel predicts better results for 70:30 of dataset.

There are certain limitations of the literature reviewed on predicting cryptocurrency prices using SVR models. A majority of studies depend on single kernels such as RBF or linear kernels [16, 17]. These kernels may not fully capture the complexity of cryptocurrency price data, which often displays both global trends and complex non-linear relationships. Some studies recognize the potential of SVR with combined kernels [2], but the specific combinations explored may not be optimal for capturing the interaction effects that are critical in cryptocurrency markets. The studies mainly concentrate on a limited range of cryptocurrencies [1, 16, 17]. The effectiveness of kernel combinations may vary based on the specific cryptocurrency and market conditions. The combined RBF polynomial and hybrid ANOVA polynomial RBF kernels provide a more data-driven approach that could potentially surpass single-kernel methods by capturing the multifaceted nature of cryptocurrency price data.

## 6. Conclusion and Future Scope

SVR with multiple kernels predicts Bitcoin (BTC), Ethereum (ETH), and Polkadot (DOT) stock prices with promising results. The robustness of SVR enabled us to introduce a multi-kernel model to capture cryptocurrency markets' intricate dynamics better. The proposed kernels with grid search offer several advantages over RBF kernels for stock price prediction in terms of flexibility, generalization ability, reduced overfitting, and ability to capture nonstationary patterns. With new kernels and hyperparameters tuned from grid search optimizations in SVR, our models are more adaptable and predictive. The fine-tuned models outperformed traditional approaches using evaluation metrics such as MAPE and $R^2$ score, proving their effectiveness. Our research found a MAPE value of 0.07772 and an $R^2$ score of 0.9999. This technique can assist investors, traders, and researchers in making informed decisions about cryptocurrencies. Adding ANOVA kernel may increase the risk of overfitting, especially with limited data. Tuning multiple hyper parameters from different kernel functions (RBF, Polynomial, ANOVA) can be challenging. It requires more computational resources to find the optimal combination. As an extension of the present research, novel kernels, improved optimization, and ensemble techniques can be used to make more comprehensive financial decisions.

## References

[1].   K. Murray, A. Rossi, D. Carraro, A. Visentin, On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep

Learning, and Ensembles. Forecasting, 5(1), (2023) 196-209. https://doi.org/10.3390/forecast5010010

[2]. H. Sebastiao, P. Godinho, Forecasting and trading cryptocurrencies with machine learning under changing market conditions. Financial Innovation, 7(3), (2021). https://doi.org/10.1186/s40854-020-00217-x

[3]. A.M. Khedr, I. Arif, M. El-Bannany, S.M Alhashmi, M. Sreedharan, Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey. Intelligent Systems in Accounting, Finance and Management, 28(1), (2021) 3-34. https://doi.org/10.1002/isaf.1488

[4]. Z. Chen, C. Li, W. Sun, Bitcoin price prediction using machine learning: An approach to sample dimension engineering. Journal of Computational and Applied Mathematics, 365, (2020) 112395.https://doi.org/10.1016/j.cam.2019.112395

[5]. D. Zhao, A. Rinaldo, C. Brookins, Cryptocurrency price prediction and trading strategies using support vector machines. arXiv: Trading and Market Microstructure, (2019). https://doi.org/10.48550/arXiv.1911.11819

[6]. M. Poongodi, A. Sharma, V. Vijayakumar, V. Bhardwaj, A. Parkash Sharma, R. Iqbal, R. Kumar, Prediction of the price of Ethereum blockchain cryptocurrency in an industrial finance system. Computers & Electrical Engineering, 81, (2020) 106527.https://doi.org/10.1016/j.compeleceng.2019.106527

[7]. Y.E.A. Seputra, Analysis of Cryptocurrency Using a Genetic Algorithm (GA). Proceedings, 83(1), (2022) 51.https://doi.org/10.3390/proceedings2022083051

[8]. J. Thomason, E. Ivwurie, Digital Assets and the Tokenization of Everything, Advancements in the New World of Web 3: A Look Toward the Decentralized Future. IGI Global, Chapter 6, (2023). https://doi.org/10.4018/978-1-6684-6658-2.ch006

[9]. F. Fang, C. Ventre, M. Basios, L. Kanthan, D. Martinez-Rego, F. Wu, L. Li, Cryptocurrency trading: a comprehensive survey. Financial Innovation 8(1), (2022) 13. https://doi.org/10.1186/s40854-021-00321-6

[10]. A. Sbrana, P.A. Lima de Castro, N-BEATS Perceiver: A Novel Approach for Robust Cryptocurrency Portfolio Forecasting. Computational Economics, (2023). https://doi.org/10.1007/s10614-023-10470-8

[11]. S. Sulistiana, M.A. Muslim, Support Vector Machine (SVM) Optimization Using Grid Search and Unigram to Improve E-Commerce Review Accuracy. Journal of Soft Computing Exploration, 1(1), (2020) 8-15.https://doi.org/10.52465/joscex.v1i1.3

[12]. F. Tang, Y. Wu, Y. Zhou, Hybridizing Grid Search and Support Vector Regression to Predict the Compressive Strength of Fly Ash Concrete. Advances in Civil Engineering, 2022(1), (2022) 3601914. https://doi.org/10.1155/2022/3601914

[13]. Y. Wu, Y. Zhou, Splitting tensile strength prediction of sustainable high-performance concrete using machine learning techniques. Environmental Science and Pollution, 29, (2022) 89198–89209. https://doi.org/10.1007/s11356-022-22048-2

[14]. S.R. Thumu, G. Nellore, Optimized Ensemble Support Vector Regression Models for Predicting Stock Prices with Multiple Kernels, Acta Informatica Pragensia, 13(1), (2024). https://doi.org/10.18267/j.aip.226

[15]. P. Xu, Q. Cao, Y. Shen, M. Chen, Y. Ding, H. Cheng, Predicting the Motion of a USV Using Support Vector Regression with Mixed Kernel Function. Journal of Marine Science and Engineering, 10(12), (2022) 1899. https://doi.org/10.3390/jmse10121899

[16]. D. Haryadi, A. Rahman, D. Marini, U. Atmaja, S. Nurgaida, Implementation of Support Vector Regression for Polkadot Cryptocurrency Price Prediction. International Journal of Informatics Visualization (JOIV), 6, (2022) 201–207. https://dx.doi.org/10.30630/joiv.6.1-2.945

[17]. S.A. Alahmari, Predicting the Price of Cryptocurrency Using Support Vector Regression Methods. Journal of Mechanics of Continua and Mathematical Sciences, 15(4), (2020) 313-322. http://dx.doi.org/10.26782/jmcms.2020.04.00023

[18]. R. Barcenas, R. Fuentes-Garcia, L. Naranjo, Mixed kernel SVR addressing Parkinson's progression from voice features, PLoS ONE, 17(10), (2022) e0275721. https://doi.org/10.1371/journal.pone.0275721

[19]. A. Aljadani, DLCP$^2$F: a DL-based cryptocurrency price prediction framework, Discover Artificial Intelligence, 2(20), (2022). https://doi.org/10.1007/s44163-022-00036-2

[20]. L.N. Mintarya, J.N. Halim, C. Angie, S. Achmad, A. Kurniawan, Machine learning approaches in stock market prediction: a systematic literature review. Procedia Computer Science, 216, (2023) 96-102. https://doi.org/10.1016/j.procs.2022.12.115

[21]. A.A. Oyedele, A.O. Ajayi, L.O. Oyedele, S.A. Bello, K.O. Jimoh, Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction. Expert Systems with Applications, 213(1c), (2023)

119233. https://doi.org/10.1016/j.eswa.2022.119233

[22]. E. Mahdi, V. Leiva, S. Mara Beh, C. Martin-Barreiro, A New Approach to Predicting Cryptocurrency Returns Based on the Gold Prices with Support Vector Machines during the COVID-19 Pandemic Using Sensor-Related Data. Sensors, 21(18), (2021) 6319. https://doi.org/10.3390/s21186319

[23]. K. Ucak, G.O. Gunel, Adaptive stable backstepping controller based on support vector regression for nonlinear systems. Engineering Applications of Artificial Intelligence, 129, (2024)107533. https://doi.org/10.1016/j.engappai.2023.107533

[24]. A. Goel, S.K. Srivastava, (2016) Role of Kernel Parameters in Performance Evaluation of SVM. Second International Conference on Computational Intelligence & Communication Technology (CICT), IEEE, Ghaziabad, India. https://doi.org/10.1109/CICT.2016.40

[25]. O. Bonita, L. Muflikhah, (2018) Comparison of Gaussian and ANOVA Kernel in Support Vector Regression for Predicting Coal Price, 2018 International Conference on Sustainable Information Engineering and Technology (SIET), Malang, Indonesia. https://doi.org/10.1109/SIET.2018.8693153

[26]. A. Mukherjee, S. Sarkar, M. Chakraborty, S. Duttagupta, A. Bhattacharya, D. Saha, P. Bhattacharya, A. Mitra, S. Gupta, Occurrence, predictors, and hazards of elevated groundwater arsenic across India through field observations and regional-scale AI-based modelling. The Science of the total environment, 759, (2021) 143511. https://doi.org/10.1016/j.scitotenv.2020.143511

[27]. Y. Jiang, H. Li, Y. Zhou, Compressive Strength Prediction of Fly Ash Concrete Using Machine Learning Techniques. Buildings, 12(5), (2022) 690. https://doi.org/10.3390/buildings12050690

[28]. I Fadil, M.Y. Helmiawan, Y. Sofiyan, Optimization Parameters Support Vector Regression using Grid Search Method, 9th International Conference on Cyber and IT Service Management (CITSM), IEEE, Bengkulu, Indonesia. https://doi.org/10.1109/CITSM52892.2021.9589028

[29]. T.T. Ngoc, C.M.T. Le Van Dai, C.M. Thuyen, Support vector regression based on grid search method of hyperparameters for load forecasting, Acta Polytechnica Hungarica, 18(2), (2021) 143-158.

[30]. The Bitcoin (BTC) dataset openly available in Yahoo finance at https://finance.yahoo.com/quote/BTC-USD/history?p=BTC-USD

[31]. The Ethereum (ETH) dataset openly available in Yahoo finance at https://finance.yahoo.com/quote/ETH-USD/history?p=ETH-USD

[32]. The Polkadot (DOT) dataset openly available in Yahoo finance at https://finance.yahoo.com/quote/DOT-USD/history?p=DOT-USD

[33]. A.J. Smola, B. Scholkopf, A tutorial on support vector regression. Statistics and Computing, 14, (2004) 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

## Authors Contribution Statement

S.R.T: Conceptualization, Data curation, Investigation, Methodology, Writing – Original draft preparation, Visualization. G.N: Supervision, Software, Validation, Writing – Reviewing and Editing.

## Funding

## Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

## Data Availability

The data that support the findings of this research are available from the corresponding author (S.R.T), based upon reasonable request.

## Has this article screened for similarity?

Yes

## About the License