Journal on Predicting Stock Market Prices Using Machine Learning

Abstract

In this study, support vector regression (SVR) is used as a machine learning technique to predict stock market prices and trends. We employ different types of windowing operators as data preprocessing techniques for SVR models. This innovative approach leverages these windowing functions to preprocess time series data, aiming to improve prediction accuracy. We applied this method to a dataset from the Dhaka Stock Exchange (DSE), focusing on the ACI group of companies. Our results demonstrate that SVR models with rectangular and flattened window operators produce reliable predictions for 1-day, 5-day, and 22-day forecasts, showing minimal error margins.

1. Introduction

The stock market is a crucial economic sector with widespread implications for investors and stakeholders. Accurate stock price forecasting is challenging due to the market's non-linear, non-parametric, and chaotic nature. Various factors, such as liquidity, human behavior, and news, influence market trends. Traditional forecasting methods include technical analysis and parametric pricing methods, but recent advancements have seen the application of machine learning techniques like neural networks (NN) and support vector machines (SVM) for predicting stock prices.

SVR, a type of SVM used for regression, focuses on minimizing the generalized error bound, which combines training error with a regularization term to control model complexity. Previous research has demonstrated the potential of SVR in financial forecasting, motivating this study's application of SVR combined with windowing functions for improved stock price prediction.

2. Methodology

This research introduces a novel approach by combining SVR with various windowing functions as preprocessing steps to select inputs for the machine learning models. We used historical stock price data from the ACI group of companies, spanning four years (2009-2012), obtained from the Dhaka Stock Exchange (DSE).

2.1 Data Preprocessing

We employed three types of windowing operators: rectangular, flatten, and de-flatten windows. These operators generate input data by selecting different window sizes and steps to feed into the SVR models. The chosen window settings for training and testing are detailed in Table 1.

Table 1: Window Settings

Windowing Operator	Model	Window Size	Step Size	Training Window Width	Test Window Width
Rectangular	All	3	1	30	30
Flatten Window	1 day	3	1	30	30
	5 days	8	1	30	30
	22 days	25	1	30	30
De-Flatten Window	All	5	1	30	30

2.2 SVR Model and Kernel Function

We selected the Radial Basis Function (RBF) kernel for its efficiency and predictive performance. The kernel parameter settings are shown in Table 2.

Table 2: Kernel Parameter Settings

Model	Kernel	C	g	ε	ε+	ε-
Model-1	RBF	10000	1	2	1	1
Model-2	RBF	10000	1	2	1	1
Model-3	RBF	10000	1	2	1	1

Three SVR models were developed to predict stock prices 1 day, 5 days, and 22 days ahead.

3. Experimental Design

The experiments were designed to test the efficacy of SVR models using different windowing operators on historical stock price data. The dataset was split into training and testing sets, and the models' performance was evaluated based on Mean Absolute Percentage Error (MAPE) between actual and predicted prices.

Model-1: 1 Day Ahead Prediction

Support Vector	Bias (b)	Weight (w) [close-2]	Weight (w) [close-1]	Weight (w) [close-0]
696	400.7	1358.9	627.1	501.1

Model-2: 5 Days Ahead Prediction

Support Vector	Bias (b)	Weight (w) [close-2]	Weight (w) [close-1]	Weight (w) [close-0]
692	381.5	825.1	734.1	-297.1

Model-3: 22 Days Ahead Prediction

Support Vector	Bias (b)	Weight (w) [close-2]	Weight (w) [close-1]	Weight (w) [close-0]
675	421.3	1719.6	1631.5	805.1

4. Results and Discussion

The performance of the SVR models with different windowing operators was assessed by comparing predicted prices with actual prices. The results showed that models using rectangular and flattened window operators had lower error margins, indicating their suitability for short-term and medium-term stock price predictions.

Performance of SVR Models

Prediction Horizon	Window Operator	Accuracy Description	MAPE
1-Day Ahead	Rectangular	High accuracy; small difference between actual and predicted prices	Low
	Flatten	High accuracy; similar to rectangular window	Low
	De-Flatten	Slightly worse than rectangular and flatten windows; still reasonably accurate	Moderate
5-Days Ahead	Rectangular	Moderate accuracy; higher MAPE compared to 1-day prediction	Moderate
	Flatten	Better performance than rectangular window; lower MAPE	Lower
	De-Flatten	Lower accuracy compared to rectangular and flatten windows; significant prediction errors	Higher
22-Days Ahead	Rectangular	Challenging to predict; higher MAPE reflects difficulty in long-term forecasting	High
	Flatten	Outperformed rectangular window; slightly lower prediction errors	Lower
	De-Flatten	Highest prediction errors; less suitable for long-term predictions	Highest

Model	Horizon	Rectangular Window	Flatten Window	De-Flatten Window
1 Day Ahead	1	0.42	0.04	7.79
5 Days Ahead	5	0.26	0.15	7.16
22 Days Ahead	22	0.22	0.22	7.61

The results indicate that the SVR models, particularly those with rectangular and flatten window operators, provide reliable short-term and medium-term stock price predictions.

5. Conclusion

This study demonstrates the potential of combining SVR with windowing operators for stock price prediction. The SVR models showed promising results, especially for 1-day and 5-day ahead predictions. Future work will focus on exploring other windowing functions and applying the model to different datasets for further validation.

6. References

Kim, K.-J., “Financial time series forecasting using support vector machines,” Neurocomputing, 2003.

Ince, H., Trafalis, T. B., “Kernel Principal Component Analysis and Support Vector Machines for Stock Price Prediction,” IEEE, 2004.

Lai, K. C., Liu, J. N. K., “Stock Forecasting Using Support Vector Machine,” International Conference on Machine Learning and Cybernetics, 2010.