

A PROJECT SYNOPSIS ON

**Forward Forecast of Stock Price Using Sliding-window  
Metaheuristic-optimized Machine Learning Regression**

TEAM MEMBERS

Akshay Purohit	1ST14CS010
Aravind B	1ST14CS023
Arun Kumar	1ST14CS025
Ashok S	1ST14CS027

## **ABSTRACT**

Time series forecasting has been widely used to determine the future prices of stock, and the analysis and modelling of finance time series importantly guide investors' decisions and trades. In addition, in a dynamic environment such as the stock market, the non-linearity of the time series is pronounced, immediately affecting the efficacy of stock price forecasts. Thus, this work proposes an intelligent time series prediction system that uses sliding-window metaheuristic optimization for the purpose of predicting the stock prices of Taiwan construction companies one step ahead. It may be of great interest to home brokers who do not possess sufficient knowledge to invest in such companies. The system has a graphical user interface and functions as a stand-alone application. The developed hybrid system exhibited outstanding prediction performance and it improves overall profit for investment performance. The proposed model is a promising predictive technique for highly non-linear time series, whose patterns are difficult to capture by traditional models.

## **INTRODUCTION**

Financial markets are highly volatile and generate huge amounts of data daily. Investment is a commitment of money or other resources to obtain benefits in the future. Stock is one type of securities. It is the most popular financial market instrument and its value changes quickly. It can be defined as a sign of capital participation by a person or an enterprise in a company or a limited liability company. The stock market provides opportunities for brokers and companies to make investments on neutral ground. Stock prices are predicted to determine the future value of companies' stock or other financial instruments that are marketed on financial exchanges. However, the stock market is characterized by nonlinearities, discontinuities, and high-frequency multi-polynomial components because it interacts with many factors such as political events, general economic conditions, and traders' expectations. Therefore, making precise predictions of stock values are challenging.

## EXISTING SYSTEM

Time series forecasting consists in a research area designed to solve various problems, mainly in the financial area. It is noteworthy that this area typically uses tools that assist in planning and making decisions to minimize investment risks. This objective is obvious when one wants to analyze financial markets and, for this reason, it is necessary to assure a good accuracy in forecasting tasks.

1. Time series forecasting consists in a research area designed to solve various problems, mainly in the financial area. It is noteworthy that this area typically uses tools that assist in planning and making decisions to minimize investment risks. This objective is obvious when one wants to analyze financial markets and, for this reason, it is necessary to assure a good accuracy in forecasting tasks.
2. Machine learning (ML) is coming into its own that can play a key in a wide range of critical applications. In machine learning, support vector machines (SVMs) have many advanced features that are reflected in their good generalization capacity and fast computation. They are also not very sensitive to assumptions about error terms and they can tolerate noise and chaotic components. Notably, SVMs are increasingly used in materials science, the design of engineering systems and financial risk prediction.
3. Support vector regression (SVR), a variant of the SVM, was developed by Vapnik et al. (1995). SVR is typically used to solve nonlinear regression problems by constructing the input-output mapping function. The least squares support vector regression (LSSVR) algorithm is a further development of SVR by Suykens (2001) and involves equality instead of inequality constraints, and works with a least squares objective function. The LSSVR approach considerably reduces computational complexity and increases efficiency compared to standard SVR.

## **DISADVANTAGES OF EXISTING SYSTEM**

- Since time series data can be formulated by regression analysis, LSSVR is very efficient when applied to the issue at hand. However, the efficacy of LSSVR strongly depends on its tuning hyperparameters, which are the regularization parameter and the kernel function. Inappropriate settings of these parameters may lead to significantly poor performance of the model. Therefore, the evaluation of such hyperparameters is a real-world optimization problem.
- Since the performance of SVR-based models strongly depends on the setting of its hyperparameters, they used to be set in advance based on the experience of practitioners, by trial-and-error, or using a grid search algorithm. Thus, finding the optimal values of regularization and kernel function parameters for SVR-based models is an important and time-consuming step. Therefore, a means of automatically finding the hyperparameters of SVR, while ensuring its generalization performance, is required.

## **METHODOLOGY**

The Firefly Algorithm (FA), which is a nature-inspired metaheuristic method, has recently performed extremely well in solving various optimization problems such as stock price forecasting and electricity price prediction. The standard FA was developed by modelling the behaviour of tropical fireflies. Notably, the smart firefly algorithm-based LSSVR has been demonstrated to be very effective in solving complex problems in civil engineering.

Recent research suggests that hybrid forecasting models can be usefully applied to the stock market's fluctuations, yielding satisfactory forecasting precision. The authors used a hybrid model to capture the linear and non-linear characteristics of a stock price time series and confirmed that hybrid forecasting models are powerful tools for practitioners in management science. A review of the literature has indicated that enhancing the effectiveness capability of least squares support vector regression based on a nature-inspired metaheuristic optimization algorithm, such as the firefly algorithm is an unsolved problem in the field of stock price prediction.

## **PROPOSED SYSTEM**

Decision to buy or sell a stock is very complicated since many factors can affect stock price. This work presents a novel approach, based on a metaheuristic firefly algorithm and least squares support vector regression (MetaFA-LSSVR), to constructing a stock price forecasting expert system, with the aim of improving forecasting accuracy. The intelligent time series prediction system that uses sliding-window metaheuristic optimization is a graphical user interface that can be run as a stand-alone application. The system makes the prediction of stock market values simpler, involving fewer computations, than that using the other method that was mentioned above.

## **ADVANTAGES OF PROPOSED SYSTEM**

- To evaluate the proposed approach, it was applied to five datasets for stocks in Taiwan, and three other stock datasets that have been used in other papers.
- Firstly, to generalize the application of the proposed system, our work uses the proposed system to estimate other stocks in similar emerging markets and mature markets, such as Vietnam, Indonesia, China, Japan, Hong Kong, Korea, Singapore, Europe, USA and India.
- Secondly, the system can be extended to analyze multivariate time series data and import raw dataset directly.
- Thirdly, profit can be maximized even when the construction corporate stock market is bullish. Finally, the development of a web-based application has been considered to improve the user-friendliness and usability of the expert system.

## SYSTEM ARCHITECTURE

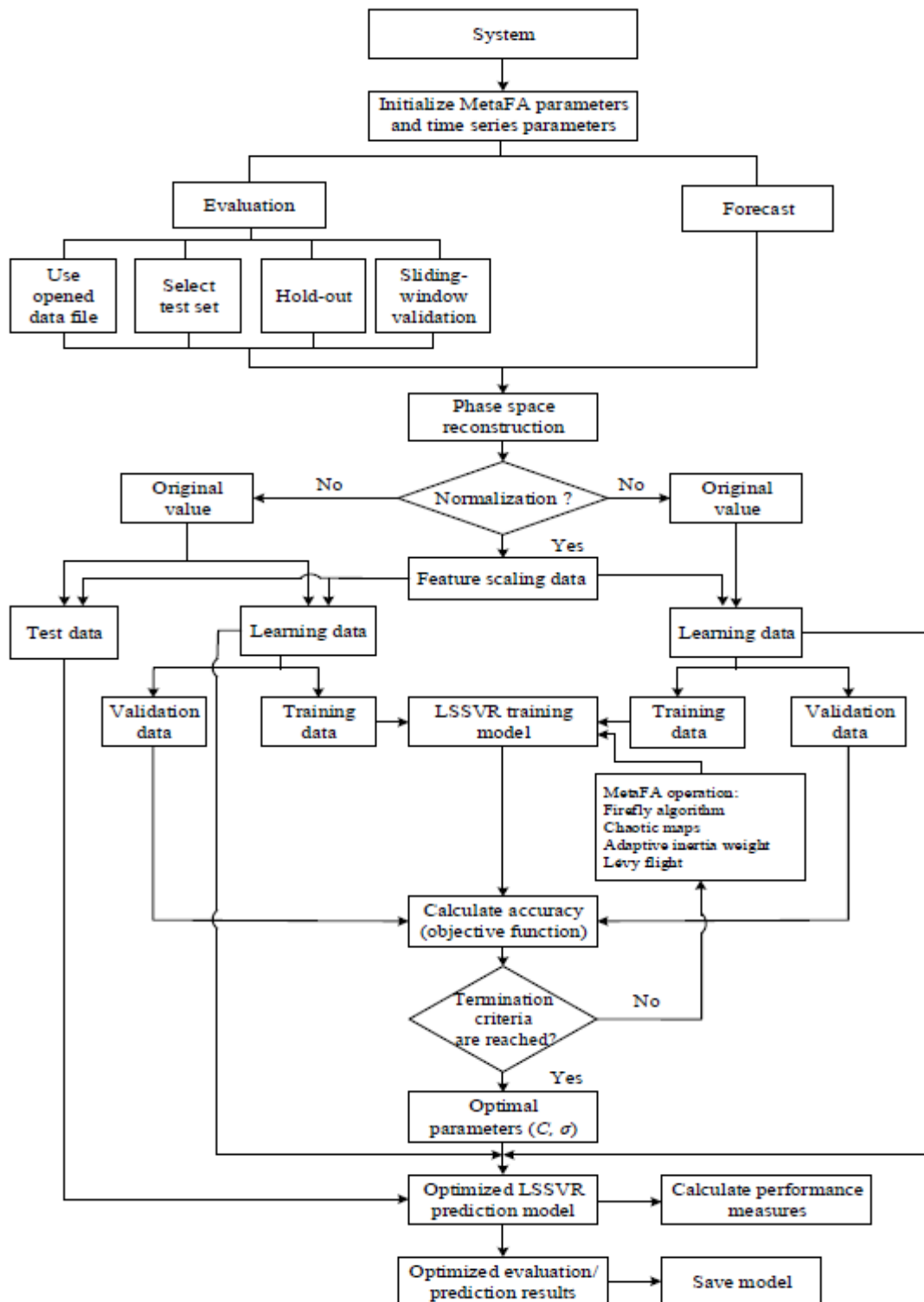


Figure 1 System Flowchart

## LIMITATIONS AND FUTURE SCOPE

The limitation of the proposed system is its computational speed, especially with respect to sliding-window validation, because of the complexity of solving large mathematical loops in the MATLAB program. The computational cost increases with the number of validations. Another weakness is the need to define many parameters of the system (MetaFA and time series parameters) though the default settings are provided. Moreover, the system does not achieve outstanding results for long-term investment – a finding that will motivate future research.

## HARDWARE REQUIREMENTS

- **System/Processor** Intel Core i3,5,7, 2.4-3.0 GHz
- **Hard Disk Space** 500 GB or more
- **RAM** 4 GB or more

## SOFTWARE REQUIREMENTS

- **Operating System** Windows/Linux
- **Coding Language** Python, Web (HTML, JavaScript...)
- **IDE/Editor** PyCharm, Brackets

\*\*\*