

# Statistical Evaluation and Prediction of Financial Time Series Using Hybrid Regression Prediction Models

Dr. M. Durairaj<sup>1\*</sup>, B. H. Krishna Mohan<sup>2</sup>

Submitted: 12/10/2021 Accepted : 03/12/2021

**Abstract:** Financial time series are chaotic by nature, which makes prediction difficult and complicated. This research employs the new hybrid model for the prediction of FTS which comprises Long Short-Term Memory (LSTM), Polynomial Regression (PR), and Chaos Theory. First of all, FTS is tested for the presence of chaos, in this hybrid model. Later, using Chaos Theory, the time series is modelled with the chaos existence. The model time series will be entered in LSTM for initial forecasts. The sequence of errors derived from LSTM forecasts is PR appropriate for error predictions. Error forecasts and original model forecasts are applied to produce the final hybrid model forecasts. Performance testing of the hybrid model (Chaos+LSTM+PR) is conducted using three categories namely foreign exchange, commodity price and stock-market indices. The hybrid model proposed in this study, in compliance with MSE, Dstat and Theil's U, is proved superior to the individual models like ARIMA, Prophet, LSTM and Chaos+LSTM. The execution of these various hybrid proposed methods is done mainly using Python, additionally, the authors used Gretl® and R for some methods respectively. Ultimately, the final result of this hybrid model describes with a better result than the existing prediction models and it is proved using various types of FTS like Foreign exchange rates, commodity prices, and stock market indices respectively. Hence, the result shows that the proposed hybrid models of Chaos+LSTM+PR achieved with better prediction rate than the existing models on the nine datasets executed.

**Keywords:** LSTM, Chaos, Polynomial Regression, Financial Time Series (FTS), Deep Learning, Time Series Prediction, Exchange rate, Stock market index, Commodity price.

This is an open access article under the CC BY-SA 4.0 license.  
(<https://creativecommons.org/licenses/by-sa/4.0/>)

## 1. Introduction

Financial Time Series is a group of financial vector (s) observations that are reported regularly, e.g. daily exchange rates, stock market values, and commodity prices, for example, are all available daily. Overall, the FTS is chaotic and transient [1]. At first, a chaotic time series is delicate, definitive, and nonlinear. [2]. Therefore, FTS is noisy, with mathematical features varying over time. The forecast is unlikely with this property [3,4]. It is often difficult to construct the proper prediction model to catch nonlinearity in the time series. As a result, forecasting FTS is a difficult and dynamic operation. In FTS prediction, the intimidating challenge is to create the appropriate predictive model to catch the subtle and apparent data changes. A variety of standard statistical techniques for the prediction of FTS, such as moving average, exponential smoothing, ARIMA, etc are contained in the literature. The predictions of the test results have not been reliable, while these techniques are statistically efficient. Subsequently, the different computational techniques suggested for the time series prediction were Swarm Intelligence-based model, Fuzzy Systems, Artificial Neural Networks (ANNs) etc., which are also seemed to offer more precise advances. However, they don't get drawbacks. The hybrid projection models in the Soft Computing paradigm, which in many permutations and combinations merge multiple

prediction models demonstrate that momentum can be gained with higher forecasts in comparison to their counterparts.

Many researchers showed that time series or hybrid time series models could work better than individual time series models [5-7]. The researchers concluded that effectively integrating multiple forecasting approaches can provide better forecasts than stand-alone methods depending on the existing research, which created the platform for presenting different hybrid time series models. Similarly, the existing study showed that the prediction accuracy of a set of several models is often needed. The combination of multiple ANNs increased the accuracy of time series forecasting, which is also reported by another study. In order to increase forecasting efficiency, reduce model variability, overcome individual model defects, and boost prediction accuracy, a hybrid model integrates two or more individual forecast models into a common model. As a result of its complexity and importance, financial time series prediction necessitates the development of more advanced and sophisticated hybrid algorithms.

Chaos theory suggested a new way of modelling a deterministic complex system's underlying non-linear dynamic behaviour by a given scalar FTS using parameters such as the lag and integration incorporated in their respective phase spaces, where lag represents time delay and integration dimension means the number of variables necessary for the nonlinear dynamics of chaotic system [8,9]. The implementation of deep neural network methods can be useful to achieve greater accuracy of prediction [10,11]. Deep learning enables ANNs to study the descriptions of data of various abstraction levels (hierarchical learning), which is part of machine

<sup>1\*</sup>Bharathidasan University, Khajamali Campus, Trichy-620023, India  
ORCID ID: 0000-0002-8666-7871

<sup>2</sup>Bharathidasan University, Khajamali Campus, Trichy-620023, India  
ORCID ID: 0000-0001-9347-1757

\* Corresponding Author Email: mohanbk28@gmail.com

learning, [12,13]. The ANNs will build a problematic and non-linear feature that maps output from the inputs. These are employed to resolve various financial challenges, such as the methods for trade execution, optimization of portfolios, stock market prediction and processing of financial intelligence [14]. This area is largely unexplored, though, and so one of the hybrid approaches like LSTM is adopted. The LSTM is a kind of Recurrent Neural Network that reads the FTS step by step [15]. It can learn long-term dependencies in a chain, as a result of which it observes non-linearity in time series as good, resulting in more reliable forecasts than linear forecasting models like ARIMA. The polynomial regression models are applied to estimate market changes and to determine the performance of various models by retaining information during the training and testing the accuracy of these predicting models applied by the known information. This research paper provides a hybrid model of FTS prediction incorporating the theory of chaos, LSTM, and PR. The FTS is primarily tested for chaos existence in this hybrid. The time series is then modelled using Chaos Theory with the chaotic existence. Then the time series is inserted in LSTM for original forecasts. PR is capable of obtaining error estimates for the error sequence derived from LSTM forecasts. To get a final forecast of the hybrid model, from the LSTM, the error estimates and original forecasts are applied.

The contributions of this paper include:

For FTS two new chaos-based hybrids prediction: Chaos with Long ShortTerm Memory and Chaos with Long ShortTerm Memory with Polynomial Regression.

Solutions to three related forecast issues include prediction of exchange rates, prediction of the stock market index and prediction of goods prices.

Comparative analysis of hybrids suggested by individual time series, including ARIMA, Prophet and LSTM prediction models.

The rest of this research paper is arranged as follows, section 2 explains the survey of the research, section 3 briefly explains the proposed method of this research, section 4 explains the datasets, tools, and model evaluation employed in this study under experimental design, section 5 explains the experimental result of the proposed method and compared with existing techniques and the final section tells the conclusion part of this research.

## 2. Literature Survey

The prediction of the FTS had been studied using various analyses or models, over the last two decades. In addition, several experiments have demonstrated in the literature that there are various appropriate approaches for effective prediction values for FTS. The research reviews the primary direction of recent features that explain the issues of predicting time series in this section.

Jonatan Henríquez et al [16] recommended a hybrid model to simulate the volatility of five US Dollar exchange rates: French-Swiss Franc, Euro, Japanese Yen, the Canadian Dollar, and British Pound. The Independent Component Analysis (ICA) was the proposed model, an example of the sequence in different components or neural networks (NNs). This approach differentiates this research from previous experiments in which ICA was used to obtain explanatory variables for estimating time-series noise or to remove time series noise. Alejandro Parot et al [17] suggested a new method to enhance the forecast precision for EUR/USD exchange-rate returns through a Vector Error Correction (VECM), post-processing, and Vector Auto-Regressive (VAR) model with the presence of ANN. The inducement should incorporate many approaches to enhance the potential of each

model to forecast. The model suggested was scored 32.5% on the baseline model that decreases RMSE by 19.3% of the better econometric model and the optimal hybrid was displayed in the result.

As its roots and improvements, Lei Ruan et al [18] suggested that the financial industry study the financial institutions' physical sales representatives, banking as an automated banking conventional, and the bank's mobile device, representing mobile financial growth in three major formats. Based on this, the authors examined three flaws in the existing mobile terminal architecture of mobile banking and suggest alternative mobile finance formats. Equipment's personalities, low carbon environment, wearability, security, offline interaction, intelligence and privacy are all "smart financial" and are embedded in five other key features. The updating of smart financial formats to satisfy customers' expectations at a larger and more profound level would result in mobile financial formats in the future. The ANN approach has its benefits in tackling the challenges of the economic environment.

In order to design future prices of gold, R. Hafezi et al [19] proposed the establishment of ANNs for the smart model. A meta-heuristic algorithm known as a Bat algorithm includes the recommended neural network to detect fluctuations. The model is compared with the study paper reported and other competitive frameworks, including the Generalized Regression Neural Networks, the Adaptive NeuroFuzzy Inference System, Multilayer Perceptron, the Radial Basis Function Network, ANN, and ARIMA. Root Mean Squared Error was an error rate to test model efficiency. Results revealed that both traditional and new forecasts were outperformed in the proposed BAT Neural Network (BNN).

During the period 1999-2019, M. Durairaj et al [20] conducted a systematic analysis of these hybrid prediction models. It is found that over these two decades there are 34 separate hybrids of FTS focused on deep learning. This complete analysis will allow financial forecasters to get an insight into current deep learning hybrids and new scientists to research deep learning in existing hybrids and to create new ones. In order to estimate stock price fluctuations, a modern long short-term hybrid memory model (SLM) integrates the LSTM model with different generalized autoregressive conditional heteroscedasticity (GARCH) models. Ha YoungKim et al [21] suggested using Index KOSPI 200, the hybrid model solutions that integrate an LSTM with a GARCH model of one or three. Further, the researchers contrast its output with current methodologies by evaluating the exponential GARCH, an exponentially weighted moving mean, GARCH, and the LSTM, as well as the DFN hybrid model integrating a DFN with a GARCH-type model. Its performance is comparable to the LSTM hybrid models proposed.

In the ModAugNet system, Yujin Baek et al [22] suggested a new method for increasing data in stock market index forecasting, which includes two modules: the LSTM avoidance overlapping module, and the LSTM prediction module. A tested outcome of the suggested model is two different representative market information (KOSPI200 and S&P500). The findings demonstrated the excellent predictability of the model proposed. Cao et al [23] suggested two models of hybrid prediction: EMD-LSTM and CEEDMAN-LSTM. The effect of noise in FTS is minimized by the decomposition of the original series by a complete ensemble empirical decomposition model with adaptive noise (CEEDMAN) and empirically mode decomposition (EMD). The denoised series are then inserted into LSTM to get predictions. The authors concluded that the hybrids proposed could provide a valid one-step forecast.

Wilfredo Tovar et al. [24] proposed a bidirectional long short-term

memory and convolutional neural network (Bi-LSTM-CNN) for the performance of the Generative Adversarial Network (GAN). It is developed to construct synthesized information that is in line with the current financial data available in order to preserve stock characteristics with positive or negative patterns to forecast potential stock trends. The newness of this proposed approach is that it implemented the idea of a composite system (Bi-LSTM-CNN) instead of a static LSTM model. The solution varies from previous solutions. Ümit Çavuş Büyükaşahin et al [25] researched the probability of integrating nonlinear and linear models to enhance forecasting. However, there are certain assumptions that these current approaches may in certain circumstances limit their efficiency. The authors have a new hybrid method as ARIMA-ANN which works in a broader context. Experimental findings indicate that the techniques to disassemble the initial data and to integrate nonlinear and linear models in the entire hybridization process are important factors in the methods' forecast efficiency.

### 3. Methodology of Financial Time Series Prediction Using Hybrid Models

In consequence of the financial rate time series prediction, the suggested hybrid strategy manages the existence of chaos. For this function, the Lyapunov exponent is used. In the scalar time series, Chaos Theory is then used to construct phase space. Optimum lag and optimum integration dimensional values are important for the creation of phase space. For selecting the optimal lag from time series, the Akaike Information Criterion (AIC) is applied. For achieving the best embedding dimension, Cao's approach is used. The phase space can be reconstructed by Chaos theory before an optimum delay and integration dimensions are reached. Ultimately, LSTM is used to obtain initial forecasts and to adjust the forecasts, the polynomial regression is applied. As a result, the hybrid model is equated with the current forecasting models like ARIMA, Prophet, LSTM, and Chaos+LSTM respectively. Figure 1 depicts the flow of the hybrid model proposed. The brief details and explanation of the hybrid models like chaos theory, LSTM, and polynomial regression are described as follows,

#### 3.1. Chaos Theory

Chaos theory is a branch of deterministic dynamics that proposes that due to the complexity of the systems involved, seemingly random occurrences can arise from ordinary equations. Chaos in time series is modelled by deriving both the embedded dimension of the associated space from the time series ( $m$ ) and the ideal lag ( $l$ ) [27]. This process transforms the given one-dimensional nonlinear time series into the corresponding multidimensional representation. A time series of scalar represented as  $Y = \{y_1, y_2, \dots, y_k, y_{k+1}, \dots, y_N\}$ . It is completely integrated into the vector  $P_j = \{y_j, y_{j+l}, y_{j+2l}, \dots, y_{j+(m-1)l}\}$  represented  $m$ -dimensional space where integration dimension is denoted as  $j = 1, 2, \dots, N - (m-1)l / \Delta t$  ( $m \geq \theta$ , where  $\theta$  denoted as attractor dimension),  $l$  represented as time delay and  $\Delta t$  denoted as the sampling time.

#### 3.1.1. Akaike Information Criterion

The Akaike Information Criterion (AIC) is applied to choose the appropriate lag when modelling the autoregressive component of data for autocorrelated time series. It defines a delay  $j$  to decrease  $\frac{\log(SSR(j))}{N} + (j+1) \frac{c(N)}{N}$  where  $SSR(j)$  is the sum of vector autoregression  $j$  lags with squared residuals, where  $n$  is the number

of observations represented as  $N$ .

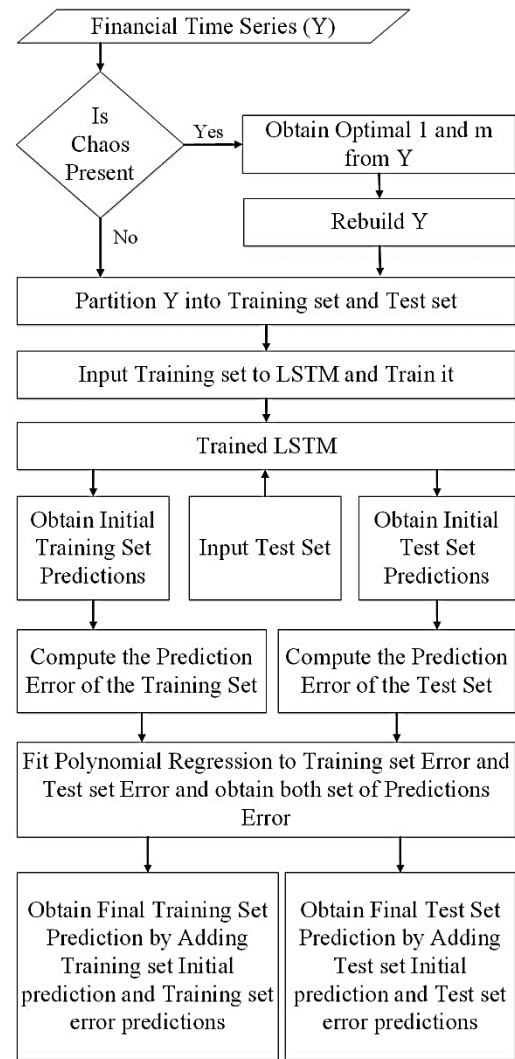


Fig 1. Proposed Approach

#### 3.1.2. Cao's method

Cao proposed to examine the minimal integration dimension for a given set. Let  $Y = \{y_1, y_2, \dots, y_k, y_{k+1}, \dots, y_N\}$  be a time series. As time delay vectors in the phase domain, it can be reconstructed as in:

$$Y_i^m = \{y_i, y_{i+1}, y_{i+2l}, \dots, y_{i+(m-1)l}\};$$

$$i = 1, 2, \dots, n - (m-1)l \quad (1)$$

Here,  $Y_i^m$  denotes the  $i^{\text{th}}$  reconstructed parameter. The nearest neighbours of the  $Y_i^m$  denoted as  $\psi Y_i^m$  are given as in

$$\psi Y_i^m = \{\psi y_i, \psi y_{i+1}, \psi y_{i+2l}, \dots, \psi y_{i+(m-1)l}\};$$

$$i = 1, 2, \dots, n - (m-1)l \quad (2)$$

It classifies equation (3),

$$\phi_2(i, m) = \frac{\|y_i^{m+1} - \psi Y_i^m\|}{\|y_i^m - \psi Y_i^m\|} \quad (3)$$

Here,  $\|\cdot\|$  is the maximum norm leading to Euclidean distance,  $y_i^{m+1}$  denotes the  $i^{\text{th}}$  parameter reconstructed, and  $\psi Y_i^m$  is its

nearest neighbour in the integration dimension  $m+1$ . They define the mean value of all  $\phi_2(i, m)$  as in equation (4),

$$E(m) = \frac{1}{n-ml} \sum_{i=1}^{n-ml} \phi_2(i, m) \quad (4)$$

$E(m)$  is only dependent on both  $l$  and  $m$ .  $E(m)$  relies exclusively on  $l$  and  $m$ . The equation (5) is defined below to examine its variation from  $m$  to  $m+1$ ,

$$E_1(m) = \frac{E(m+1)}{E(m)} \quad (5)$$

If some reason  $E_1(m)$  stops changing  $m \geq m_0$  for some  $m_0$ , then  $m_0$  is the necessary minimum embedding dimension.

### 3.2. Long Short-Term Memory

Issues in stock prices and indexing are predicted by an LSTM model. An LSTM has three gates: an input gate to determine whether the new input is to be entered, a forget gate that deletes information that is not important, and the output gate that specifies which output information. These three gates are analog gates based on the sigmoid function which works on the range 0 to 1.  $h_t$  is the hidden output layer sequence at time  $t$ ,  $x_t$  is the series of input layers at time  $t$ , and  $A$  is the hidden memory cell layer. In the process of data transfer in the hidden layer,  $C_t$  is special information that contains the state of the cell. The main objective of the transmission is to change the cell state and output efficiency  $h_t$  at times  $t$ . Figure 2 shows the particular transition in the LSTM cell. The neural network layer in each yellow block reads data from  $h_{t-1}$  and  $x_t$ . Cross blocks are calculative units that are designed for vector operation.

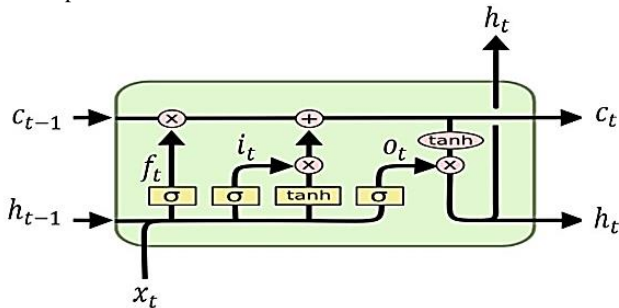


Fig 2. LSTM structure

The structure of LSTM is displayed in the above figure 2, and the following are the equations of the front of an LSTM unit with a forget gate [28]:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (6)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (7)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (8)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (9)$$

$$h_t = o_t \odot \sigma_h(C_t) \quad (10)$$

Where  $f_t$  denotes as forget gate, input gate indicated as  $i_t$ , output gate is represented as  $o_t$ ,  $C_t$  is denoted as a cell state,  $h_t$  is represented as hidden state,  $\sigma$  denoted as activation function, and

the Hadamard product is represented as the operator named as  $\odot$ .  $W$ ,  $U$ , and  $b$  are the weight matrices trained by the author. In order to understand the more complex features of the results, the author trained these models using the stacked LSTM. Stacked LSTM is commonly used for sequential data tasks like predictions of FTS, speech recognition, and anomaly detection.

### 3.3. Polynomial Regression

Polynomial models are applied where the correlation between the analysis and describing variables is curvilinear. Sometimes, polynomials may also model a nonlinear interaction in a limited set of explanatory variables. The trend model is based on the polynomial order and the coefficient value. In continuous-time, a polynomial of degree  $N$  can be written as follows:

$$y(t) = \sum_{i=0}^N c(i)(nT)^i \quad (11)$$

For a uniformly sampled discrete-time, the continuous-time,  $t$ , is represented as  $t = nT$ , where  $n$  is an integer and  $T$  is the sampling period.

#### 3.3.1. Quadratic Polynomial Regression

Quadratic polynomial and an AR time-series are carried out for the exploration of data representation. In equation (11),  $N$  is equal to 2 for any quadratic polynomial, as it shows the second-degree polynomial regression [29]. Thus,

$$y(nT) = c(0) + c(1)(nT) + c(2)(nT)^2 \quad (12)$$

$$y(nT - T) = c(0) + c(1)(nT - T) + c(2)(nT - T)^2 \quad (13)$$

$$y(nT - 2T) = c(0) + c(1)(nT - 2T) + c(2)(nT - 2T)^2 \quad (14)$$

Using equations (12) and (13), it will be

$$y(nT) = y(nT - T) + c(1)T + 2c(2)nT^2 - c(2)T^2 \quad (15)$$

Subsequently, using equations (13) and (14), it will be

$$y(nT - T) = y(nT - 2T) + c(1)T + 2c(2)nT^2 - 3c(2)T^2 \quad (16)$$

Now, using equations (15) and (16), which provides

$$y(nT) = y(nT - T) + [y(nT - T) - y(nT - 2T)] + 2c(2)T^2 \quad (17)$$

Therefore,

$$y(nT) = 2y(nT - T) - y(nT - 2T) + 2c(2)T^2 \quad (18)$$

Equation (18) may be written as  $y(n) = 2y(n-1) - y(n-2) + 2c(2)T^2$  by eliminating  $T$  from the indices. It is evident that  $q = 2$ ,  $a(1) = 2$ ,  $a(2) = -1$ , and  $\mu = 2c(2)T^2$ .

Therefore, the following can be concluded:

Every second-degree i.e., quadratic polynomial may be perfectly described by an AR(2) time series.

In time-series analysis, every quadratic polynomial will have the same coefficient values, i.e.,  $a(1) = 2$  and  $a(2) = -1$ .

The  $\mu = 2c(2)T^2$  represents the time-series constant term.

This indicates that every quadratic polynomial with various values of  $c(0)$  and  $c(1)$  but the same value of  $c(2)$  has the same AR(2) representation, i.e., through the same values of  $a(1)$ ,  $a(2)$ , and  $\mu$ .

### 3.4. Proposed Course of Action

The proposed hybrid approach is described as follows. In FTS forecasts it is very important to minimize the average forecast error. Let  $Y = \{y_1, y_2, \dots, y_k, y_{k+1}, \dots, y_N\}$  be a time series

with  $N$  observations recorded at times  $t = \{1, 2, 3, \dots, k, k+1, \dots, N\}$ . Then perform the following:

- Check  $Y$  to see if chaos exists. If chaos exists, the  $Y$  dimension of integration ( $m$ ) can achieve the optimal delay ( $l$ ).

- Upon obtaining optimal delay and integration dimensional values, reconstruct phase space from  $Y$ .

- Partition step space after it has been rebuilt  $Y$  into  $Y_{Train} = \{y_1; t = lm + 1, lm + 2, \dots, k\}$  and

$$Y_{Test} = \{y_1; t = k + 1, k + 2, \dots, N\}$$

Train LSTM by intake  $Y_{Train}$  to LSTM and get initial predictions for the training set using the equation (19).

$$\begin{aligned} \dot{y}_t &= f_1(y_{t-1}, y_{t-2l}, y_{t-3l}, \dots, y_{t-ml}) \\ t &= lm + 1, lm + 2, \dots, k \end{aligned} \quad (19)$$

Obtain initial test set predictions based on the input  $Y_{Test}$  to trained LSTM by substituting  $t = k + 1, k + 2, \dots, N$  in equation (19). Replace  $t = k + 1, k + 2, \dots, N$  in the following equation, to compute the training set and test set of prediction errors.

$$\begin{aligned} e_t &= y_t - \dot{y}_t \\ t &= lm + 1, lm + 2, \dots, k \end{aligned} \quad (20)$$

Adjust Polynomial Regression (PR) to a training set of errors and estimate the training set error by the following equation. Similarly, adjust PR to test set of errors and get forecasts for test set errors by substituting  $t = k + 1, k + 2, \dots, N$  in the following equation.

$$\begin{aligned} \dot{e}_t &= f_2(e_t) \\ t &= lm + 1, lm + 2, \dots, k \end{aligned} \quad (21)$$

- Using the equations below, add training set initial predictions and predictions of training set error to get final training set predictions. Add the initial test set predictions and the test set error prediction to achieve the predictions of the final test set by substituting  $t = k + 1, k + 2, \dots, N$  in the following equation.

$$\begin{aligned} \ddot{y}_t &= \dot{y}_t + \dot{e}_t \\ t &= lm + 1, lm + 2, \dots, k \end{aligned} \quad (22)$$

## 4. Experimental Design

The experimental design of this study is described with the various types of datasets used for the process of the proposed method, performance of tasks for the model and tools used for the method and the model evaluation carried out for the performance of the proposed method.

### 4.1. Datasets for the Study

The efficiency of proposed hybrids is observed with the use of various datasets utilized in this paper. These regular data sets include:

From the Federal Reserve, the three foreign exchange rates are collected which are Indian Rupee/United State Dollar, Singapore Dollar/United State Dollar, and Japanese Yen/United State Dollar.

- Three stock market indexes are compiled from the investment websites: Shanghai Composite Index, Standard & Poor 500, and Nifty 50.

- Three commodity prices are compiled from Investing website, namely Soya Beans Price, Crude Oil Price, and Gold Price.

**Table 1.** Datasets Used for the Study

Dataset	Dates	Number of Observations	Training Set	Testing Set
INR/USD	01-Jan-2014 to 01-Jan-2020	7527	6022	1505
JPY/USD	01-Jan-2014 to 01-Jan-2020	7535	6028	1507
SGD/USD	01-Jan-2014 to 01-Jan-2020	7534	6028	1506
S&P 500	01-Jan-2014 to 01-Jan-2020	5695	4556	1139
Nifty 50	01-Jan-2014 to 01-Jan-2020	6009	4808	1201
Shanghai Composite	01-Jan-2014 to 01-Jan-2020	7099	5680	1419
Crude Oil Price	01-Jan-2014 to 01-Jan-2020	7605	6084	1521
Gold Price	01-Jan-2014 to 01-Jan-2020	7620	6096	1524
Soya Beans Price	01-Jan-2014 to 01-Jan-2020	7755	6204	1551

The datasets along with the appropriate dates, several remarks, a training set, and a test set are given in Table 1. The FTS problem forecast is modelled as a supervised learning problem. Thus, each dataset is divided into a training set (80%) and a test set (20%). The chaos that exists in each dataset is verified for all of these datasets and it is found. After that, the space of phase of each data set is reconstituted with a corresponding optimum delay and integration dimension.

### 4.2. Performance of Tasks and Tools Used

The analysis of the proposed method prediction (Chaos+LSTM+PR) is performed in the system requirements with specifications of Intel Core i3 @ 3.5GHz processor and 6GB-DDR3 RAM. In addition, the hybrid methods of this study are done using Python, Gretl<sup>®</sup> and R, and the details are mentioned below.

1. Chaos Existence used for Nold's package and function is `lyap_r()` with Python.
2. For Optimum Delay, the function is used AIC with Gretl<sup>®</sup>.
3. For Optimum Integration Dimension used for `nonlinearTseries` and function is used `estimateEmbeddingDim(.)` with R Tool.
4. Importing Data, the package used Pandas, the function used `read_csv(.)` with Python.
5. Partitioning Data, the package used `sci-kit-learn`, the function used `train_test_split(.)` with Python.
6. Fitting ARIMA to data, `mModule` used `statsmodels`, the function is `ARIMA(.)` with Python.
7. Prophet to data, Module used `fbprophet`, the class used `Prophet` with python.
8. LSTM to data, the package used `keras`, the function used `LSTM()`, `Predict()` with Python.
9. PR to data, the module used `sci-kit-learn`, the function used `PolynomialFeatures(.)` with Python.
10. MSE Computation, module used `sci-kit-learn`, function used `mean_squared_error(.)` Python.
11. Dstat Computation performed with Python.
12. Theil's U Computation performed with Python.
13. Statistical Significance, Package used to forecast, Function used `dm.test(.)` with R.

Various activities are conducted throughout this experimentation. The best p, d and q values of the ARIMA model are acquired from the pmdarima module of python using Auto Arima(.). Additionally, the estimate Embedding Dim(.) method from the nonlinear Time-series package implemented Cao's method described above.

### 4.3. Model Evaluation

Three separate target functions, including MSE, Dstat, and Theil's U, to name a few obtain the optimal solution, where both the MSE value and the standard deviation is minimum. Theil's Inequality Coefficient, Mean Squared Error, and Directional Change Statistics have been taken into account to calculate the efficiency of this proposed model. These indices have checked the accuracy and uncertainty of this hybrid model. The authors returned both of these measurements on time to analyze the time-series properties of errors in forecasts [30].

#### 4.3.1. Mean Squared Error

The Mean Squared Error (MSE) calculates how precisely the model forecasts the result by the calculation of the average squares of errors. The test set estimates the value of the next term after the model has been compiled by estimating the model parameters. Then the outcomes of the results are translated to the origin from differential form. The analysis tests both the train set and the test set by MSE. It is helpful when researchers are involved with significant errors that have a relatively bigger negative effect over the smaller equivalent [38]. An MSE value nearer 0 indicates statistical suitability.

$$MSE = \frac{\sum_{t=1}^N (y_t - \hat{y}_t)^2}{N} \quad (23)$$

Where N is the series of numbers and the forecasted value are observed at the time step t. Comparative analyses are shown in the result section.

#### 4.3.2. Directional Change Statistic

Dstat shall calculate the directional change of time series. The tests for assessment include the estimation of the proper matching number of the actual and predicted values,  $y_t$  and  $\hat{y}_t$  respectively, in the testing set with respect to the directional change is concerned. Statistics of signs can be expressed as,

$$Dstat = \frac{1}{N} \sum_{t=1}^N a_t * 100\% \quad (24)$$

$$a_t = \begin{cases} 1, & \text{if } (y_{t+1} - y_t) * (\hat{y}_{t+1} - \hat{y}_t) \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

The higher the Dstat value, the better movements of time series are discovered by the model. The standardized mean square errors only calculate prediction in terms of levels so that these statistics are desirable. Thus, the predictive efficiency can be calculated by the accuracy of the gradient (Dstat) forecasts and the changes in the sign.

#### 4.3.3. Theil's Inequality Coefficient

Theil's Inequality coefficient referred to be U, determines how much closer a forecast time series is to the actual time series. The U value generally ranges from 0 to 1. A value of zero is combined with a perfect prediction, (e.g.)  $y_t = \hat{y}_t$  meaning that U=0 fits well with all observations. Additionally, a value of one is associated with a forecast that on average has the same error as a "naive" no change forecast (i.e.) U=1 indicates the performance as low.

$$U = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t)^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}_t)^2}} \quad (25)$$

Many of these performances' corresponding equations measurements, at the time (t) the actual and estimated values are denoted as  $y_t$  and  $\hat{y}_t$ , and the number of predicted values are represented as N.

## 5. Experimentation and Results Discussion

During experiments with different conditions, the data sets are used together. The LSTM architecture here used includes a completely integrated 50-nodes dense layer with each node being enabled by the ReLU activation feature. The adam optimizer is used as a loss function for the LSTM for 500 epochs. LSTM, Chaos+LSTM and Chaos+LSTM+PR are entered by scaled values using MinMaxScaler. During PR modelling errors, The method used is second-degree polynomial regression. Each dataset's results are listed here.

### 5.1. INR/USD

The results of the test set of INR/USD are represented in Table 2. Compared to Standard LSTM according to MSE and Theil's U, Chaos+LSTM could not predict well. The ARIMA model could, according to MSE and Theil's U, outperform a prophet. Standard LSTM could outperform ARIMA and Prophet in accordance with MSE and Theil's U.

**Table 2.** The results of the INR/USD test set

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA(3,1,1)	6.592027	49.83	0.0007361
Prophet	13.45557	52.49	0.0016148
LSTM	0.115719	49.36	1.32020e-05
Chaos+LSTM (l=8, m=10)	2.361804	51.49	0.0002748
Chaos+LSTM+PR (l=8, m=10)	<b>2.2717e-08</b>	<b>99.93</b>	<b>2.6010e-12</b>

Therefore, from Table 2, the proposed hybrid is certainly superior to independent comparison versions in respect to the evaluation criteria. It shows the accuracy of predictions is better with the actual values which are depicted in Figure 3.



**Fig 3.** Prediction of INR/USD Testset using the Hybrid Model

Figure (3) represents the prediction of the test set of the IND/USD using the hybrid model. This figure shows forecasted values and actual values indicate that Chaos+LSTM+PR could provide even tighter forecasts such that the curve of the expected predictions

overlapped with the curves of real values. The prediction rate is increased according to the date wise and it reached over a maximum of 74. The prediction rate is increased according to the date wise; the prediction rate starts from the rate of 62 in Jan 2014 and reached over a maximum of 74 in earlier to Jan 2021. The proposed hybrid method has predicted a low rate of 58, between Jan 2014 and Jan 2015, and the final prediction of a rate of 71 approximately is obtained in Jan 2021.

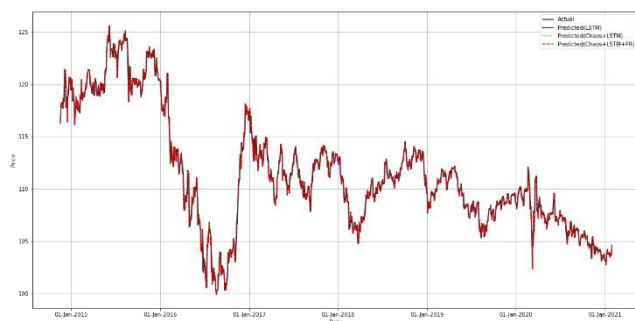
### 5.2. JPY/USD

The results of the JPY/USD test set are shown in table 3. It is clear from the table that the hybrid model gives a better result than the actual prediction techniques of individual forecasting models with the performance of measures of the evaluation parameters.

**Table 3.** The results of the JPY/USD test set

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (0,1,0)	221.758833	49.10	0.010098
Prophet	779.334350	47.97	0.040086
LSTM	0.37960545	50.43	1.5309 e-05
Chaos+LSTM (l=1, m=10)	0.40028878	50.29	1.6183e-05
Chaos+LSTM+PR (l=1, m=10)	<b>2.7287 e-06</b>	<b>99.93</b>	<b>1.1014 e-10</b>

The table shows the accuracy of the predictions is better with the comparison of the actual values shown in Figure (4). Nevertheless, according to MSE and Theil's U, Chaos+LSTM could not predict well compared to standard LSTM.



**Fig 4.** Prediction of JPY/USD Testset using the Hybrid Model

Figure (4) represents the prediction of the test set of the JPY/USD using the hybrid model. This figure concludes that the predicted values using the forecasting models of Chaos+LSTM+PR results with a slightly increased rate related to the actual values with the date-wise prediction. Using the proposed method, the predicted rate has started from Jan 2014 with approx. value of 105. This test set prediction achieved over the maximum of 125 rates in the starting and it steps down to a rate value of 100. Moreover, the rate has reached 118 in earlier to Jan 2017 and with a rate of 109 in Jan 2021, the rate drops marginally.

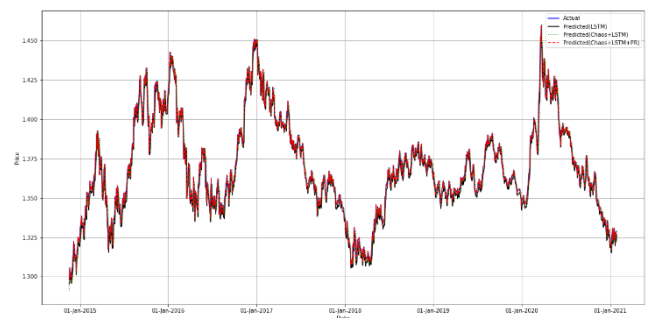
### 5.3. SGD/USD

The results of the SGD/USD test set are shown in Table 4. From the table, the measurements of MSE, Dstat, and Theil's U shows that the actual forecasting models of ARIMA, Prophet, LSTM, Chaos+LSTM are lower than the proposed model of Chaos+LSTM+PR.

**Table 4.** The results of the SGD/USD test set

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (1,1,0)	0.0335077	50.92	0.0103374
Prophet	0.0898944	46.94	0.0302199
LSTM	2.0250 e-05	49.40	5.5359 e-06
Chaos+LSTM (l=10, m=8)	0.00023338	49.33	6.3459 e-05
Chaos+LSTM+PR (l=10, m=8)	<b>1.3553 e-11</b>	<b>99.93</b>	<b>3.7029 e-12</b>

The shows the predictions are strongly correlated in comparison to the actual values shown in Figure (5). However, when compared to MSE's Standard LSTM and Theil's U, Chaos+LSTM could not predict well.



**Fig 5.** Prediction of SGD/USD Testset using the Hybrid Model

Figure (5) represents the prediction of a test set of the SGD/USD by employing the proposed hybrid model. The proposed model of Chaos+LSTM+PR prediction achieved to get a slightly higher rate of values than the actual models. Using the proposed hybrid method, the prediction has started from Jan 2014 with approx. value of 1.27 and end to Jan 2021 with the rate of 1.34. The predicted rate occurred to get the maximum rate of 1.45 in Jan 2017, after it reached the average rate and the minimum rate of 1.22 between Jan 2014 and Jan 2015.

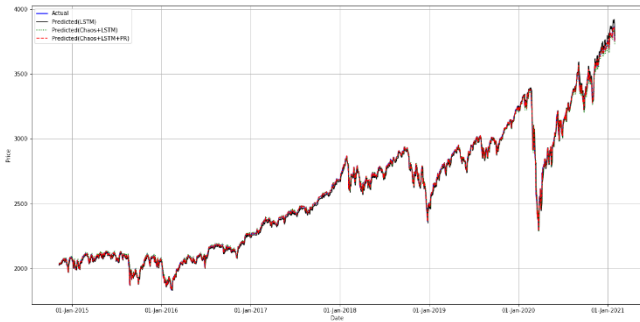
### 5.4. S&P 500 Stock Index

The results of the S&P 500 Stock Index test set are shown in Table 5. From the table, the study results with the present models of ARIMA, prophet, LSTM, and Chaos+LSTM are much lower compared to the Chaos+LSTM+PR with the parameters of evaluation of the result.

**Table 5.** Results of the S&P 500 Stock Index

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA	285444.117	55.13	0.026778
Prophet	25983.4741	49.95	0.001980
LSTM	490.933425	49.25	3.8276e-05
Chaos+LSTM (l=7, m=8)	32678.6553	50.65	0.0027027
Chaos+LSTM+PR (l=7, m=8)	<b>0.06774960</b>	<b>99.91</b>	<b>5.2714e-09</b>

The table displays the accuracy, in contrast to actual values, forecasted values are higher as shown in Figure (6). However, in comparison to LSTM with the performance parameters, Chaos+LSTM did not forecast well.



**Fig 6.** The Hybrid Model was used to predict the S&P 500 Stock Index Testset.

The hybrid model forecasts of the test set of the S&P 500 stock index are shown in Figure (6). The illustration depicted above describes that compared with the actual values, the predicted proposed methods value gives better performance with slightly higher values. The rate of 3200 values is acquired with the maximum rate of values by the prediction of proposed models.

### 5.5. Nifty 50 Stock Index

The results of the Nifty 50 Stock Index test set are shown in Table 6. It is clear from the table that the Chaos+LSTM+PR model is the most likely to get higher results than the present models of ARIMA, prophet, LSTM, and Chaos+LSTM. Figure (7) represents the actual values which are accurately low compared to the forecasted values.

**Table 6.** The test set of the Nifty 50 Stock Index yielded the following results.

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (0,1,1)	838809.710	52.70	0.0043867
Prophet	1716969.596	50.62	0.01008946
LSTM	6572.658455	51.54	3.45740e-05
Chaos+LSTM (l=7,m=8)	1140903.572	49.20	0.00552445
Chaos+LSTM+PR (l=7,m=8)	<b>0.600146522</b>	<b>99.91</b>	<b>3.15453e-09</b>

However, in relation to the LSTM norm in compliance with the performance measures, Chaos+LSTM could not easily forecast. In addition, the nifty 50 stock index compares the proposed model results with the Covid pandemic results and the stock market predictions.

Figure (7) represents that the hybrid model predictions of the test set of the Nifty 50 Stock Index. Using the proposed hybrid models, the prediction price starts over Jan 2015 with the approximate price of 8900 and ended in Jan 2021 with the predicted approximate price of 12200. The price rate in nifty 50 stock increases the same as the increasing rate of S&P 500. Similarly, the predicted price in nifty 50 has a stepdown value of 7000 between Jan 2016 and Jan 2017.

### 5.6. 5.6 Shanghai Composite Index

The results of the test set of the Shanghai Composite Stock Index is represented in Table 7. The resultant table described that the proposed model of Chaos+LSTM+PR results is comparatively equal to the current forecasting models with the compliance of evaluation parameters. The predicted values accuracy is better with the hybrid models suggested in this study compared to the individual models.



**Fig 7.** Prediction of Nifty 50 Stock Index Testset using the Hybrid Model

**Table 7.** The Shanghai Composite Index test set's results

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (3,1,3)	933325.909	54.19	0.06433685
Prophet	2474342.012	48.13	0.20496063
LSTM	2603.222918	49.18	0.00013619
Chaos+LSTM (l=8,m=7)	36964.73470	51.23	0.00195842
Chaos+LSTM+PR (l=8,m=7)	<b>0.109618890</b>	<b>99.92</b>	<b>5.76047 e-09</b>

However, in comparison with the normal LSTM according to the parameters of the evaluation, the Chaos+LSTM cannot be well forecasted.



**Fig 8.** Prediction of Shanghai Composite Index Testset using the Hybrid Model

Figure (8) represents that the predictions of the test set of the Shanghai Composite Index by applying the hybrid model proposed. The test set predictions of the shanghai composite index give slightly increased results with the proposed models than the current models. The prediction of price in the shanghai composite index using the proposed hybrid models starts over Jan 2014 with the price of 2000, and the price increases dramatically to the approximate rate of 5200 between Jan 2015 and Jan 2016. After that, the price drops to 3500 in Jan 2016. Then the price is predicted as constant around 3500 to 2500 till Jan 2021.

### 5.7. Crude Oil Price (USD)

Table 8 shows the outcomes of the test set of Crude Oil Prices in accordance with USD. From the table, the hybrid model (Chaos+LSTM+PR) in conjunction with the performance measures is very much superior to the comparison models of the individuals.



**Table 8.** The results of the Crude Oil Price (USD)

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA(3,1,2)	2953.539167	51.21	0.18946605
Prophet	4525.145236	52.13	0.24282701
LSTM	1.3635525	48.78	0.00018750
Chaos+LSTM (l=6,m=9)	8.8273485	49.63	0.00121101
Chaos+LSTM+PR (l=6,m=9)	<b>3.39347e-08</b>	<b>99.93</b>	<b>4.65323e-12</b>

Table 8 indicates forecasts accurately better with the real values. Thus, in relation to the LSTM norm in compliance with the performance measures, Chaos+LSTM could not easily forecast. Figure (9) represents that the hybrid model's predictions of the test set of Crude Oil prices (USD). The results of the test set of crude oil price (USD) predictions are the slightly lower rate of values by the present models compared with the proposed models. The prediction price for crude oil (USD) using the proposed hybrid model starts over Jan 2014 with the maximum price of 100 approximately, after that, the price drops down over Jan 2016 with the minimum predicted price of 25 approximately. Then the predicted price continues with the constant price of around 40 to 75 approximately till Jan 2021.



**Fig 9.** Prediction of Crude Oil Price (USD) Testset using the Hybrid Model

**5.8. 5.8 Gold Price (USD)**

The results of the gold price test set in US dollars are represented in Table 9. The table describes that the proposed method results better than the present independent models with the presence of parameters of the evaluated results.

**Table 9.** The results of the Gold Price (USD)

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (2,1,1)	92948.1638	49.93	0.02150899
Prophet	501419.446	51.96	0.08654559
LSTM	272.959089	46.58	7.83139e-05
Chaos+LSTM (l=10,m=8)	1828.57969	52.29	0.00051785
Chaos+LSTM+PR (l=10,m=8)	<b>4.11548e-05</b>	<b>99.93</b>	<b>1.18221 e-11</b>

Thus, this table represents that the hybrid models suggested in this research obtain an accurately better result than the comparison of individual models. The predicted values using the hybrid models perform with accurate results. Figure (10) displays the hybrid model predictions of the test set of Gold prices (USD). The gold price forecasts in the test set (USD)

displayed with the sine waveform of a graph in the above figure, also shows that the hybrid model results are comparatively related to the actual result. The prediction of price for gold using the proposed hybrid models starts over Jan 2014 with the predicted price of 1450 approximately and ends in Jan 2021 with the price of 1600 approximately. The graphical representation shows the gold price like a zigzag waveform. The minimum price was predicted in Jan 2016 with 1050 approximately and the maximum price was predicted between Jan 2019 and Jan 2021 with 1560 approximately.



**Fig 10.** Prediction of Gold Price (USD) Testset using the Hybrid Model

**5.8.1. Soya Beans Price (USD)**

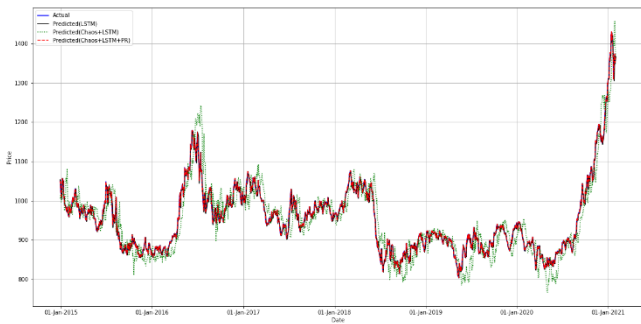
The price of soyabeans in US dollars based on the results of the test set are represented in Table 10. The table shows that the Chaos+LSTM+PR of the proposed model occurs comparatively equal to the present model of MSE, Dstat, and Theil's U. assessment of ARIMA, Prophet, LSTM, and Chaos+LSTM. Consequently, the individual comparison models help for the comparison results of the hybrid model proposed presents the better solution to forecast the soya beans price with accurate values.

Figure (11) represents that the hybrid model predictions of the test set of Soyabeans Price (USD). The above figure displays the estimated values of forecasting Models are compared to real values, revealing that Chaos+LSTM+PR can produce very similar predictions. The predicted values of both the present and the proposed method results are displayed from higher to lower rate date-wise. Using the proposed hybrid model, the prediction of soyabeans price (USD) starts from Jan 2014 with the price of 1320, afterwards the price increases marginally over 1520 between Jan 2014 and Jan 2015. Then the price drops over 900 in earlier to Jan 2015, after that the price maintains constant around 1180 to 800 with some ups and downs from Jan 2015 to Jan 2021.

**Table 10.** The Price of Soya Beans results in a Test Set (USD)

Model for Forecasting	Mean Squared Error	Directional Change Statistic	Theil's Inequality Coefficient
ARIMA (0,1,0)	208672.937	50.35	0.06975013
Prophet	687001.670	49.19	0.16405291
LSTM	173.007352	49.58	8.607095e-05
Chaos+LSTM (l=10,m=8)	1718.10239	51.25	0.00085458
Chaos+LSTM+PR (l=10,m=8)	<b>0.00010501</b>	<b>99.93</b>	<b>5.243315e-11</b>

As shown in Table 10 the values forecasted by applying the hybrid model proposed are superior as compared to the actual values forecasted using the existing individual models.



**Fig 11.** Prediction of Soyabeans Price (USD) Testset using the Hybrid Model

## 6. Conclusions

FTS predictions are useful for real investing and for analyzing the general direction of other financial market indexes. FTS is noisy and its statistical characteristics differ over time. The prediction is not possible with this characteristic. It's still difficult to develop the correct prediction model, which can observe the non-linearity of the time series. In this article, a new hybrid model was presented Chaos+LSTM+PR, which will address the issue of FTS prediction. Initially, the FTS is tested for the prevalence of chaos in this hybrid. Then Chaos Theory was modelled the chaos present in the time series. For the optimum time series lag, the AIC is used. For achieving optimum measurements of integration, Cao's approach is used. Using the chaos theory, the phase space is reconstructed until the optimum delay and the optimum integration dimension are obtained from the time series. The model time series was entered in LSTM for initial forecasts. The LSTM prediction error sequence is suitable for PR to achieve error predictions. To modify the predictions, polynomial regression is applied. In order to achieve final forecasts on the hybrid model, error predictions and original LSTM predictions are applied. Three different categories of FTS are used as foreign currencies, asset values and stock market indexes to measure the effectiveness of the hybrid model. The proposed hybrid with MSE, Dstat and Theil's U output is higher than the corresponding models listed in these sections. The hybrid suggested can also be applied for different financial and non-FTS applications. Furthermore, the proposed method was achieved to get a better result than the present method using nifty 50. The execution of analysis of the hybrid model was done using python, Gretl®, and R. The python was mainly covered for many methods respectively. Finally, the comparison of the existing models with the hybrid model is calculated using Diebold and Mariano test. It is evident from this test that Chaos+LSTM+PR is more successful in forecasting all dataset models. Hence, the final result of the proposed approach explains with a better result than the current forecast models and it is proven using various forms of FTS including foreign exchange rates, stock prices, and stock market indices respectively.

## References

- [1] S. Alonso-Monsalve, A.L. Suárez-Cetrulo, A. Cervantes, and D. Quintana, "Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators." *Expert Systems with Applications*, vol.149, pp.113250, 2020.
- [2] Z. Shu, P.W. Chan, Q. Li, Y. He, and B. Yan, "Characterization of daily rainfall variability in Hong Kong: a nonlinear dynamic perspective." *International Journal of Climatology*, 2020.
- [3] T.T. Souza, and T. Aste, "Predicting future stock market structure

- by combining social and financial network information." *Physica A: Statistical Mechanics and its Applications*, vol.535, pp.122343, 2019.
- [4] E. Blankespoor, "The impact of information processing costs on firm disclosure choice: Evidence from the XBRL mandate." *Journal of Accounting Research*, vol.57, no.4, pp.919-967, 2019.
- [5] L. Fioramonti, L. Coscieme, and L.F. Mortensen, "From gross domestic product to wellbeing: How alternative indicators can help connect the new economy with the Sustainable Development Goals." *The Anthropocene Review*, vol.6, no.3, pp.207-222, 2019.
- [6] L. Meligkotsidou, E. Panopoulou, I.D. Vrontos, and S.D.Vrontos, "Quantile forecast combinations in realized volatility prediction." *Journal of the Operational Research Society*, vol.70, no.10, pp.1720-1733, 2019.
- [7] S. Smyl, "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting." *International Journal of Forecasting*, vol.36, no.1, pp.75-85, 2020.
- [8] Y. Lu, R. Luo, and Y. Zou, "Morphological Analysis for Three-Dimensional Chaotic Delay Neural Networks." *Journal of Mathematics*, 2020.
- [9] T.A. Fathima, V. Nedumpozhimana, Y.H. Lee, S.Winkler, and S. Dev, "A Chaotic Approach on Solar Irradiance Forecasting." *In IEEE, Photonics & Electromagnetics Research Symposium-Fall (PIERS-Fall)*, pp. 2724-2728, 2019.
- [10] G. Kumar, S. Jain, and U.P. Singh, "Stock market forecasting using computational intelligence: A survey." *Archives of Computational Methods in Engineering*, pp.1-33, 2020.
- [11] W. Bao, J. Yue, and Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory." *PloS one*, vol.12, no.7, pp. e0180944, 2017.
- [12] T. Niu, J. Wang, H. Lu, W. Yang, and P. Du, "Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting." *Expert Systems with Applications*, vol.148, pp.113237, 2020.
- [13] S.I. Lee, and S.J. Yoo, "Multimodal deep learning for finance: integrating and forecasting international stock markets." *The Journal of Supercomputing*, vol.76, pp.1-19, 2019.
- [14] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, "Stock Market Prediction Based on Generative Adversarial Network." *Procedia computer science*, vol.147, pp. 400-406, 2019.
- [15] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction." *IEEE Communications Magazine*, vol.57, no.6, pp.114-119, 2019.
- [16] J. Henriquez, and W. Kristjanpoller, "A combined Independent Component Analysis-Neural Network model for forecasting exchange rate variation." *Applied Soft Computing*, vol.83, pp.105654, 2019.
- [17] A. Parot, K. Michell, and W.D. Kristjanpoller, "Using Artificial Neural Networks to forecast Exchange Rate, including VAR-VECM residual analysis and prediction linear combination." *Intelligent Systems in Accounting, Finance, and Management*, vol.26, no.1, pp.3-15, 2019.
- [18] L. Ruan, C. Li, Y. Zhang, and H. Wang, "Soft computing model-based financial aware spatiotemporal social network analysis and visualization for smart cities." *Computers, environment, and urban systems*, vol.77, pp.101268, 2019.
- [19] R. Hafezi, and A. Akhavan, "Forecasting gold price changes: Application of an equipped artificial neural network." *AUT Journal of Modeling and Simulation*, vol.50, no.1, pp.71-82, 2018.
- [20] M. Durairaj, and B.H. Krishna Mohan, "A Review of Two Decades of Deep Learning Hybrids for Financial Time Series Prediction." *International Journal on Emerging Technologies*, vol.10, no.3,

pp.324–331, 2019.

- [21] H. Y. Kim, and C. H. Won, “Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models.” *Expert Systems with Applications*, vol.103, pp. 25-37, 2018.
- [22] Y. Baek, and H. Y. Kim, “ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module.” *Expert Systems with Applications*, vol.113, pp.457-480, 2018.
- [23] J. Cao, Z. Li, and J. Li, “Financial time series forecasting model based on CEEMDAN and LSTM.” *Physica A: Statistical Mechanics and its Applications*, vol.519, pp. 127-139, 2019.
- [24] W. Tovar, “Deep Learning Based on Generative Adversarial and Convolutional Neural Networks for Financial Time Series Predictions.” arXiv preprint, 2020. arXiv:2008.08041
- [25] Ü.Ç. Büyüksahin, and Ş. Ertekin, “Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition.” *Neurocomputing*, vol.361, pp.151-163, 2019.
- [26] M. Durairaj, B.H. Krishna Mohan, “Stock Marketdata Analysis Using Random Forest and Deep Learning-RNN Models.” *Engineering Skill Development*, vol.14, no.8, 2021.
- [27] V. Ravi, D. Pradeepkumar, and K. Deb, “Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms.” *Swarm and Evolutionary Computation*, vol.36, pp.136-149, 2017.
- [28] S. Kim, and M. Kang, “Financial series prediction using Attention LSTM.” arXiv preprint arXiv:1902.10877, 2019.
- [29] A.K. Nandi, “Data Modeling with Polynomial Representations and Autoregressive Time-Series Representations, and Their Connections.” *Ieee Access*, vol. 8, pp.110412-110424, 2020.
- [30] D. Pradeepkumar, and V. Ravi, “Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network.” *Applied Soft Computing*, vol.58, pp.35-52, 2017.