COMPARATIVE ANALYSIS OF SELECTED TIME SERIES FORECASTING APPROACHES FOR INDIAN MARKETS

Ankit Tripathi[1], Arpit Tripathi[1], Oldřich Trenz[1], Pawan Kumar Mishra[1]

[1]Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic

ABSTRACT

Financial market analysis and prediction have been topics of interest to traders and investors for decades. This study assesses the performance of selected time series prediction methods like deep learning algorithms (Long short-term memory model (LSTM)), traditional statistical models (Seasonal Auto Regressive Integrated Moving Approach with eXogenous regressors (SARIMAX)), and advanced ensemble learning algorithms (XGBoost and FB-Prophet) using real-world data from the Indian financial market. The stock prices of Reliance Company serve as a case study, enabling a thorough evaluation of predictive accuracy and errors of the models. A pre-processing approach has been proposed and implemented, integrating significant economic factors (Gold Price, USD to INR conversion, Consumer Price Index (CPI), Wholesale Price Index (WPI) and Indian 10-year yield bond) and evaluated with technical metrics (Mean squared error, Mean Absolute Error and R2 Score). The study investigates how the inclusion of these factors impacts prediction accuracy across the selected time series prediction methods. The comparative evaluation of models before and after the pre-processing method sheds light on the evolving predictive accuracy of LSTM, SARIMAX, FB-Prophet, and XGBoost. The study showed that the SARIMAX (extension of ARIMA with seasonality and exogenous factors) and XGBOOST performed relatively well with the proposed approach while LSTM and FB prophet (though advanced) did not perform as expected in Indian financial markets. This research contributes to advancing the understanding of time series forecasting in the financial market of India, offering practical insights for decision-makers and researchers.

Keywords: Financial Time Series, Stock Market Prediction, Deep Learning in Finance, Ensemble Learning in Economics, ARIMA and XGBoost Analysis

JEL Classification: C22, C53, G17

https://doi.org/10.11118/978-80-7509-990-7-0167

1 INTRODUCTION

Predicting the trends of financial markets has always been a challenge for people from various fields such as investors, academic researchers, financial analysts, traders, and brokers. Despite professionals using different strategies for short-term investment, long-term investment, or sheer academic pursuit, forecasting results play a vital role. In the studies of Obthong et al. in 2020 and Sonkavde et al. in 2023, financial markets comprised different segments such as bond markets, derivates markets, stock markets and commodity markets (Obthong et al., 2020; Sonkavde et al., 2023). A wide variety of techniques have been used in the past by traders such as swing trading, event-driven trading, momentum trading, trend-following trading, scalping, position trading and state-of-the-art algorithmic trading. However, all these trading methods have inherent limitations due to the nature of stocks. The inherent non-parametric, nonlinear, and non-stationary nature of stock prices, coupled with the noise in time series data, adds complexity to the prediction task (Abu-Mostafa and Atiya, 1996; Farid et al., 2023). Furthermore, this coupled with the influences from the political and economic environment around the stock market gives it further complexity to forecast the price. As traditional statistical methods, including fundamental and technical analyses, grapple with limitations such as lagging indicators and prediction inaccuracies, the spotlight has shifted to cutting-edge machine learning and deep learning models in algorithmic trading. These artificial intelligence models based on historical observations of variable forms a relationship which is further used to extrapolate the data as its future predictions. The calculations and estimations previously done via statistics are being done by methods of machine learning such as the Long Short Term Memory method (LSTM) (Chen et al., 2015), Auto Regressive integrated Moving Approach (ARIMA) (Khashei and Bijari, 2011), SARIMA (Seasonal Auto-Regressive integrated Moving Approach) (Lee et al., 2008), Holt-Winters Exponential Smoothing (Dassanayake et al., 2021), FB-Prophet (Sharma et al., 2022) introduced by Facebook and XGBoost (Gumelar et al., 2020). While some of these methods have been around since the 1970s, others have been developed in the past decade with many hybrid models having the foundation of these basic models.

It is also important to know that majority of models known works on Artificial neural networks (ANN) as it is one of the most important type of non-parametric, non-linear time series model which have been used for time series forecasting (Khashei and Bijari, 2011). While the implementation of these models had been examined on various markets such as Hong Kong Stock Markets, Shanghai Stock Index, A-share market stocks in China, S&P 500 Future, Shenzhen Component Index (SZI) and Chinese Securities Index (CSI) but a relatively smaller number of research has been done on Indian markets (Liu et al., 2023, 2024; Nasiri and Ebadzadeh, 2023). The studies by authors Sapre, Gori and Seah in 2023 quotes that "India is the fifth largest economy in the world and is 3rd in purchasing power parity according to International Monetary fund (IMF)" (Sapre and Gori, 2023; Seah, 2022). This can also be interpreted as there is an immense potential in the Indian market that investors, traders, and brokers can take advantage of and multiply their investment several times. But to get maximum returns, it is necessary to have more data regarding the performance of various algorithmic models in Indian markets, the factors influencing the markets and its segments, the relation between these factors and performance of models and results of integration of different factors in different segments of markets. To gain insights into the methods, approaches or techniques that have been used or are currently being developed for stock price forecasting especially in the context of Indian market, studies such as Mehtab et al. in 2021 which focuses on utilizing LSTM-based models for stock price prediction, Singh and Borah in 2014, concentrating on forecasting State Bank of India (SBI) stock index prices through particle swarm optimization and fuzzy time series (FTS-PSO), and Mehtab and Sen in 2021, centred on predictions using Convolutional Neural Networks (CNN) and LSTMs, provide valuable insights in understanding the background and current trends in literature (Mehtab et al., 2021; Singh and Borah, 2014). The findings show that majority of publications have focused their research on development of new hybrid models and their application in different domains. In the domain of financial markets, especially stock price forecasting, majority of the studies have also focused in establishing the relation between macroeconomic variables and stock prices but have not focused on the change in performance it brings to these forecasting models (Dhingra and Kapil, 2021; Subburayan et al., 2021). The methodologies are also majorly focused on the implementation of a single approach than the comparative evaluation of multiple models.

Another relevant thing to note is that the time span, nature of economy and sector (macroeconomic variables and industry type) affects the nature of data used for algorithm training. Therefore, to the best of author's knowledge, this study, provides the missing knowledge for Indian markets and reliance industry for the decade of 2nd December 2013 to 30th November 2023. This studies focuses on comparative analysis of selected models (LSTM, SARIMAX, XGBoost and FB-prophet) AND introducing the preprocessing approach of integration of macroeconomic variables and assessing the impact on selected models AND the implementation of data on Reliance Industry Ltd. (2nd December 2013 to 30th November 2023) which is a multinational company holding stocks for different businesses. This combination gives the study a unique perspective and offers the insights for newer age models to be built in Indian markets or segments.

2 CURRENT STATE

Investing in stock markets has been one of the approaches that has been employed by various traders, investors, businesses and dealers over the past decades to earn a significant sum of money over a period. In the past, fundamental techniques like candle charts analysis, analysing earning reports, dividend yield, quantitative analysis using statistics or mathematics model were used to forecast the prices and get better investment benefit ratio. After the implementation of Regulation National Market System in 2005, people started to focus on machine learning models due to its advantages which is speed, anonymity and certainty of execution with predefined criteria and strategies which can also be seen from the rise in number of publications in 2005 till date (Litzenberger et al., 2012). This has come today in form of various methods being developed to predict the stock prices considering all the possibilities.

The implementation of basics models like LSTM, ARIMA, GRU and others to various stock markets have provided the foundation to newer models. This has provided the data required in terms of how the market shifts when facing a crisis, or when a government change, or when a new economic policy is implemented. This data is the very precursor with which researchers try to compare the newly hybrid models. When looking with a cursory point of view, most of the models have been deployed or tested on American markets, Chinese markets, or European markets but not many models are tested in Indian markets (Liu et al., 2023, 2024; Nasiri and Ebadzadeh, 2023). India being the world fifth largest economy in terms of GDP in 2023 still does not have the amount of data, there should have been. When searched with Scopus with the query "time series forecasting in India" refined with the keyword "financial market" only 137 documents come to the view with being the study on 'Bombay Stock Exchange', 'CNX Nifty' and 'S and P' as oldest using vanilla GARCH model (Karmakar, 2005). The number of search results can easily be interpreted as 'low' compared to the larger economies in the world. This study focuses on filling that literature gap and provide basic supporting evidence that will encourage the state of art hybrid model to be tested and deployed in Indian stock markets.

The study has been focused on implementation of machine learning models mainly LSTM, ARIMA, FB Prophet and XG-Boost on Indian markets. The study by Kumaria et al. in 2023, that have focuses on the LSTM to analyse how the neural network anticipates to stock's price (Kumaria et al., 2023). In the study by Chatterjee et al. in 2022, GARCH based hybrid models are employed to analyse the impact of these models on banking, IT and pharma sector

(Chatterjee et al., 2022). In the study by Srivastava et al. in 2023. The author used NIFTY 50 dataset along with financial and social indicators to predict the stock price with LSTM, SVM, KNN, Random Forest and gradient boosting regressor models (Srivastava et al., 2023). Another study by J.P.S.Kumar et al. in 2023 compares the ARIMA and RNN-LSTM on Sensex and nifty dataset (J.P.S. Kumar et al., 2023). The study by Sharma et al. in 2022, employed the Fb-prophet model in comparison to ARIMA on NSE and BSE dataset (Sharma et al., 2022). Most of the studies have utilised either one or two basic models but not four including FB prophet and comparing them to evaluate the results. The evaluation metrics (MSE, MAE and R2 score) has also been same to evaluate the efficiency of models throughout various studies, making them the priority in this study. Another differentiating factor is that most of these studies have employed the models on basic features of any stock price data (Open, high, low, close, adj. close and date) but have not included the major economic factors or variables.

The studies by Baranidharan et al. in 2021, Dhingra and Kapil in 2021 and Hussain et al. in 2012 have supported the relationship between stock prices exchanges and macroeconomic variables but have not implemented them to the selected models and have not performed a comparative analysis (Dhingra and Kapil, 2021; Hussain et al., 2012; Subburayan et al., 2021). The study by Agarwal et al. in 2019, have utilised the technical indicators to support the hypothesis using O-LSTM model only (Agrawal et al., 2019).

The factor that these economic factors have been utilised with the integration of four different models make our study distinguished.

2.1 Selection of Macro-economic factors

Economic indicators, macroeconomic variable or interest rates play a significant role in influencing the financial markets especially stock prices. The economic factors chosen for integration are gold prices (Gold futures historical prices, 2023), Indian Rupees to United States Dollar (USD INR historical data, 2023), Indian 10-Year Bond yield price (India 10-year Bond Historical Data, 2023), Wholesale Price Index (India wholesale price index (WPI), 2023) and Consumer Price Index (India consumer price index (CPI), 2023).

Over the years, a lot of studies have indicated the relationship between gold prices and stock prices. In the study by Smith in 2001, the authors have provided the empirical evidence between gold price and stock prices by using four gold prices and six stock prices for the United States indices over the period 1991 to 2001 (Smith, 2001). The correlation was also unveiled in India by Bhunia in 2012, in which they which took stock prices of the National Stock Exchange (NIFTY) and the gold prices over April 2001 to March 2011 and by Patel in 2013, in which they took the Gold prices, Sensex , BSE 100 and Nifty for January 1991 to December, 2011 and performed ADF Test (Dickey and Fuller, 1981), Johansen's cointegration test (Johansen, 1995) and Granger Causality test (Granger, 1969) to find the underlying pattern in the relationship (Bhunia and Das, 2012; Patel, 2013). Thus, we have used this indicator to provide more information about the market to the models and to increase their efficiency.

The exchange rate has also been chosen to be one of the indicators as the various literatures have focused on the forex returns and stock prices. In the study by Batra et al. in 2020, the author concluded a negative relationship between the variables through Granger Causality test by using the data for 20 years i.e. from 25th January 2000 to 25th January 2020 (Batra et al., 2020). In the study by Lakshmanasamy in 2021, the author concluded that there is a positive relationship between BSE SENSEX return and Euro/Rupee and there is a negative relationship between USD/INR, GBP/INR and the stock prices (Lakshmanasamy, 2021). Also, the prices of Gold in the dataset are in USD and after the incorporation of the price conversion, the analysis of result became much simpler.

The Indian 10-Year Bond yield price is a relatively new factor is incorporated in the study to see the impact in the stock prices forecasting results. In the study by Panigrahi in 2022, the authors have quoted that it is an economic indicator which acts as a benchmark in equity market too (Panigrahi, 2022). They have also focused to determine a long-term relationship between stock market and long-term interest rate (Indian 10-year bond yield) using the Johansen's cointegration test and vector error correction mechanism. In the study 'Identifying significant macroeconomic Indicators for Indian Stock Markets' by Aithal et al. in 2019, the authors have taken 44 indicators among which the variable 'government bond 10-year yield' is an indicator and they have used this with the features of NSE Nifty and BSE SENSEX indices and have performed a regression analysis (Aithal et al., 2019). Thus, the factor becomes highly important to assess the impact of economic variables on stock market.

The other factors are WPI (Wholesale Price Index) and Consumer Price Index (CPI) which have been added to the economic variables list. The Wholesale Price Index usually represents the change in the price of goods sold by wholesale across India and consumer price index represents the change in price of goods or services from the point of view of consumers. They both are crucial factors impacting the interest rate or currency rates and other macroeconomic variables which in turn impact the stock prices. Both of these indicators are used to measure inflation rate in India. In the studies by Hussain et al. in 2012 and Raheem Ahmed et al. in 2017, they have tried to assess the change in KSE 100 index through the variables that includes CPI and WPI and concluded that CPI and returns have a long-term relationship (Hussain et al., 2012; Raheem Ahmed et al., 2017).

3 METHODS

3.1 Selecting the models

As the domain of time series forecasting is significantly vast, various models have been used already for the implementation of stock price forecasting with newer models being developed every day. Thus, choosing any four models will be followed by the question why was that model selected. The bibliometrics offers a solution for this problem with the help of various analyses through quantifiable metrics as depicted in Figure 1. This helps in understanding

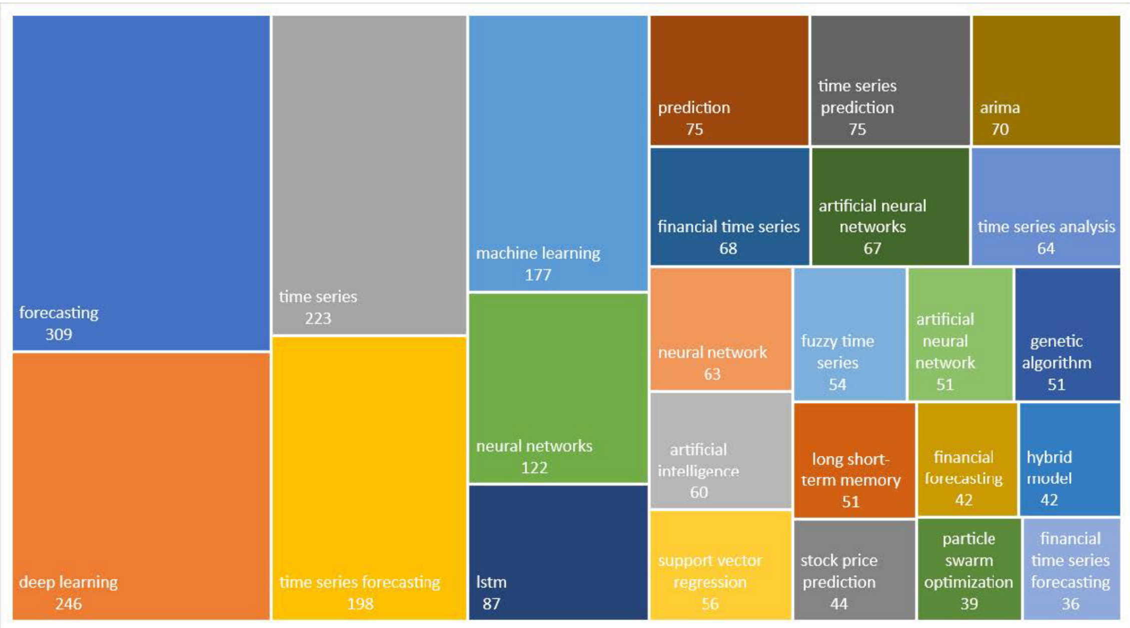

Fig. 1: Tree map of most frequent keywords.
Source: compiled by the author

that with respect to time series forecasting, financial markets and artificial intelligence, the most used models or approaches are LSTM and ARIMA.

Based on it, some models have been selected among which, two are traditional or foundational that have been used the highest number of times or have highest number of occurrence and two which are relatively state of art models. The selected state of art model is ensemble learning models like XG-Boost that also had a relation with time series forecasting. Another model like FB prophet was selected based on literatures. This will also be used to fill the literature gap in the domain and to understand the newer age model that have been classified as hybrid models. Thus, the models selected for the comparative analysis were LSTM, SARIMAX, FB-Prophet, XG-Boost.

3.2 Database and Search Strategy

While analyzing multiple time series could provide more comprehensive findings by encompassing a broader spectrum of the market, this study focuses on Reliance Industries Limited due to its prominence and influence in the Indian stock market. Reliance Industries is not only one of the largest companies in India by market capitalization but also plays a crucial role in various sectors, including energy, petrochemicals, textiles, natural resources, retail, and telecommunications. Its diverse business operations and significant market presence make it an ideal candidate for understanding broader market trends and economic influences (Biswas, 2018). By focusing on a key player like Reliance Industries, insights that are likely reflective of larger market behaviors and economic conditions can be gained, while maintaining a manageable scope for detailed analysis. The database chosen to extract the information about the stock prices of Reliance Industries Limited was Yahoo Finance (Reliance Industries Limited,2023). The domain offers various features including the historical data which have been used from December 02, 2013, to November 30, 2023. This data has been pre-processed and extracted in an CSV file with the daily frequency and as shown in Table 1.

3.3 Data Cleaning and Preprocessing

The necessary libraries were imported for the implementation of selected models. Standard Scalar has been to scale the values and to make model more efficient. Sequential, Adam (optimiser) and tensor flow keras have been imported to build and train the LSTM model. For Fb-Prophet, 'prophet' which is facebook's time series forecasting library for time series forecasting. For XGBoost, 'XGBoost' module which is used for creating and training the model. The

Date	Open	High	Low	Close	Adj Close	Volume	Gold Price	USD INR conversion	Indian 10 year bond price	CPI	WPI
02-12-2013	422.4952	427.4235	422.1485	423.6344	397.6064	3332837	1221.9	62.325	9.048	0.1124	7.52%
03-12-2013	423.4858	428.6865	421.3312	426.9034	400.6745	4238703	1220.8	62.37	9.069	0.1124	7.52%
04-12-2013	426.4081	429.6771	422.5942	424.0059	397.955	4043235	1247.2	62.06	9.09	0.1124	7.52%
05-12-2013	431.906	434.3825	429.9495	431.4106	404.9049	5784625	1231.9	61.763	9.108	0.1124	7.52%
06-12-2013	430.9897	433.2185	426.5567	429.2809	402.9059	3369441	1229	61.435	9.165	0.1124	7.52%

Tab. 1 Dataset merged with economical indicators
Source: compiled by the author

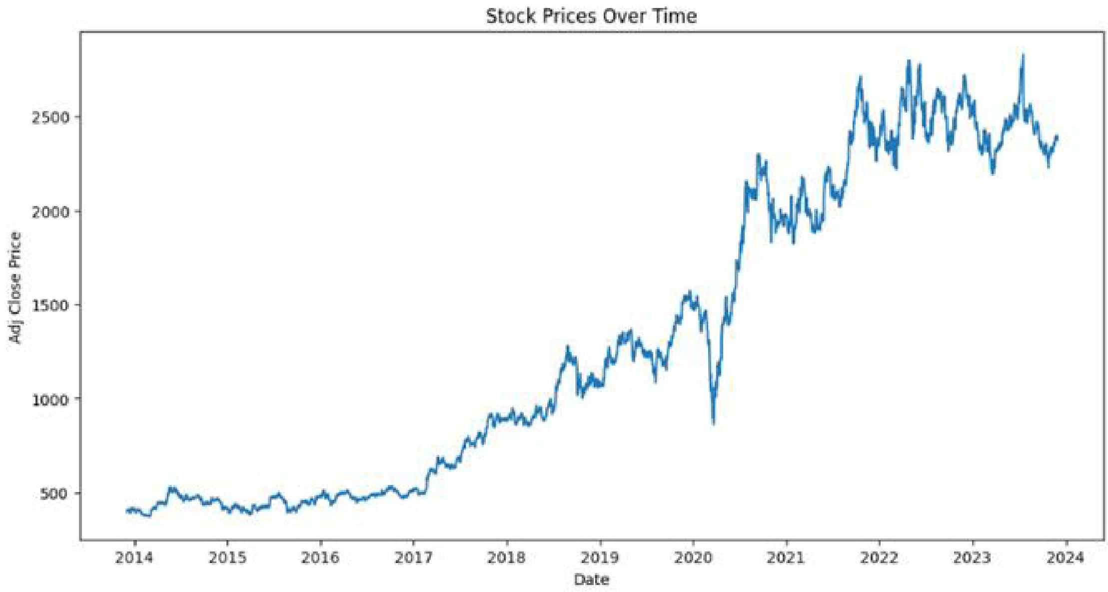

Fig. 2: Plotting thes tock prices
Source: compiled by the author

data that have been extracted in the 'csv' file was imported in the google colaboratory to work. *Adjusted Close price* was chosen with work with as it reflects the price of stock after inclusion of other factors. Some basics steps of data preprocessing were implemented to prepare the data for building the models and to make it more efficient such as:

1. Conversion of date to datetime format and setting 'Date' as index.
2. Dropping the rows with missing values.
3. Dropping the null values and dealing with NaN values.
4. The variables with high correlation coefficient were selected.
5. In case of Fb-prophet, Renaming the columns like date: 'ds' and Adj Close as 'y' to be compatible with Prophet model

After the merging of macro-economic variables:

1. As the WPI had '%' after values, so removing the percentage sign form WPI column and converting it to float.
2. Extracting the features and target and normalizing the features using Standard Scalar.

The selected evaluation metrics were Mean squared error, Mean absolute error and R2 score.

3.4 Implementation of machine learning algorithms

After the basic steps of data cleaning and preprocessing, the visualisation of stock prices was plotted with '*Date on x-axis*' and '*Adjusted close price on y-axis*' as depicted in Figure 2.

The dataset was split into two parts i.e., features (independent variables) and target (dependent variable). A standard scaler was applied to normalize the features by removing the mean and scaling to unit variance, ensuring that all features contribute equally to the model. For the models (LSTM, SARIMAX, XG-Boost and FB-prophet), training-testing splits were employed. 80% of the data was allocated for training, with 20% reserved for testing, following a standard 80:20 ratio. This approach was kept common for all the selected models to make comparative analysis general and efficient.

For LSTM, the sequential model was built using neural networks with input as features.shape (Obthong at al., 2020) which had 5 features. An input layer was provided with the number of input features. Two dense layers with 64 and 32 neurons were provided each followed by a dropout layer with a dropout rate of 20% to prevent overfitting. The activation function used was Rectified Linear Unit (ReLU) due to its ability to mitigate the vanishing gradient problem, its computational efficiency, and its tendency to converge faster compared to other activation functions like sigmoid or tanh. The output layer had a single neuron to predict the target variable. The model was compiled with the 'Mean Squared Error (MSE)' loss function, the 'Adam' optimizer (learning rate of 0.001), and 'Mean Absolute Error (MAE)' as the evaluation metric. The model was trained for 100 epochs. The forecasting plot is shown in Figure 3a). The study have also used an alternate extension of ARIMA i.e., SARIMAX (Seasonal Autoregressive Integrated Moving Average with eXogenous Factors). This type of model can take other variables into account while predicting the stock prices.

The best SARIMAX order was found to be (0,1,2) and no seasonal order was chosen. The autocorrelation and partial autocorrelation function (ACF and PACF) graphs were also plotted to determine order in the data. Augmented dickey-fuller test was also used to ensure staionarity in the data and to determine appropriate diffrencing order‘d’ but as the datasize was small

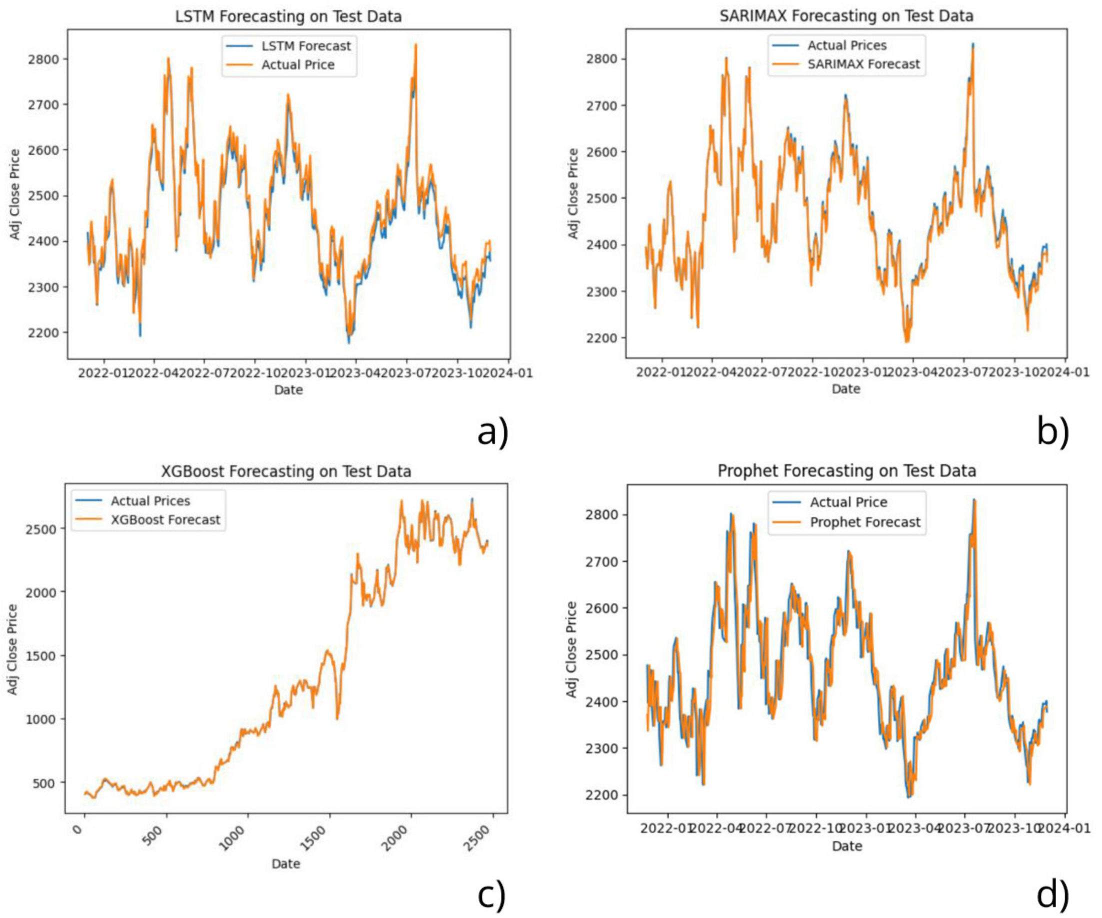

Fig. 3: a) Forecasting results by LSTM. b) Forecasting plot for SARIMAX. c) Forecasting plot for XGBoost. d) Forecasting plot for FB prophet.

Source: compiled by the author using (Google Colaboratory- Fb prophet before integration with economic variables, 2023.; Google Colaboratory- LSTM before integration with economic variables, 2023.; Google Colaboratory- SARIMAX before integration with economic variables, 2023.; Google Colaboratory- XGBoost before integration with economic variables, 2023.

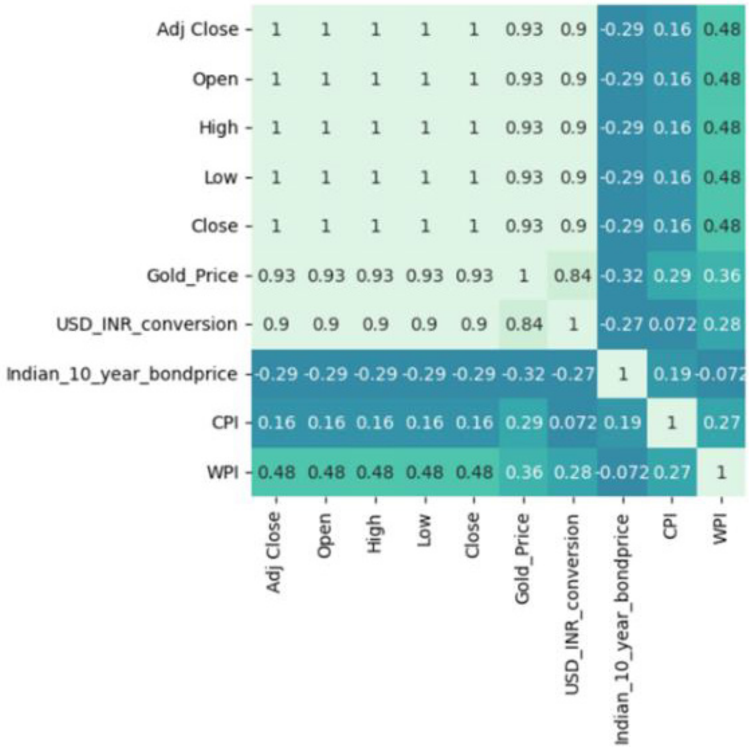

Fig. 4: Figure4: Correlation matrix.
Source: compiled by the author using Google Colaboratory- XGBoost after integration with economic variables, 2023

thus to prevent overfitting th best order was not chosen initally. The model was run with various orders to obtain the best result. The exogenous variables were the other features like open, high, close, and low. The evaluation metrics were root mean squared error, mean absolute error and r squared score. The results of forecasting on test dataset are given in Figure 3b).

For XGBoost, the model was prepared using XGBRegressor (extreme Gradient Boosting) from the XGboost library, which has been used for optimized and efficient implementation, high-speed and performance. The hyperparameters involved were learning rate which controls the contribution of each tree and was set to deafult (0.1), max_depth which represents the maximum depth of a tree and the default of 6 was used, n_estimators which represents number of trees to be fit (deault value of 100) and subsample fraction of samples to be used for fiiting the individual trees (default value of 1). The plot has been shown in Figure 3c).

For FB prophet, the model was initialised with daily seasonality. The other hyperparameters like yearly seasonality, seasonality mode and changepoint prior scale were set to default. The prediction of the model has been showed in Figure 3d).

3.5 Implementation of machine learning algorithms on pre-processed dataset merged with economic variables

The preparation and processing of the dataset (consisting of stock prices and economic variables) followed similar approach of pre-processing earlier. The basic pre-processing steps were taken from the starting like to drop the null values and renaming of the columns to avoid conflict between features. The values provided in the dataset of WPI and CPI were monthly rather than daily, thus we had to duplicate the value throughout the month to ensure smooth application of dataset on models. Figure 4 shows the correlation among all variables

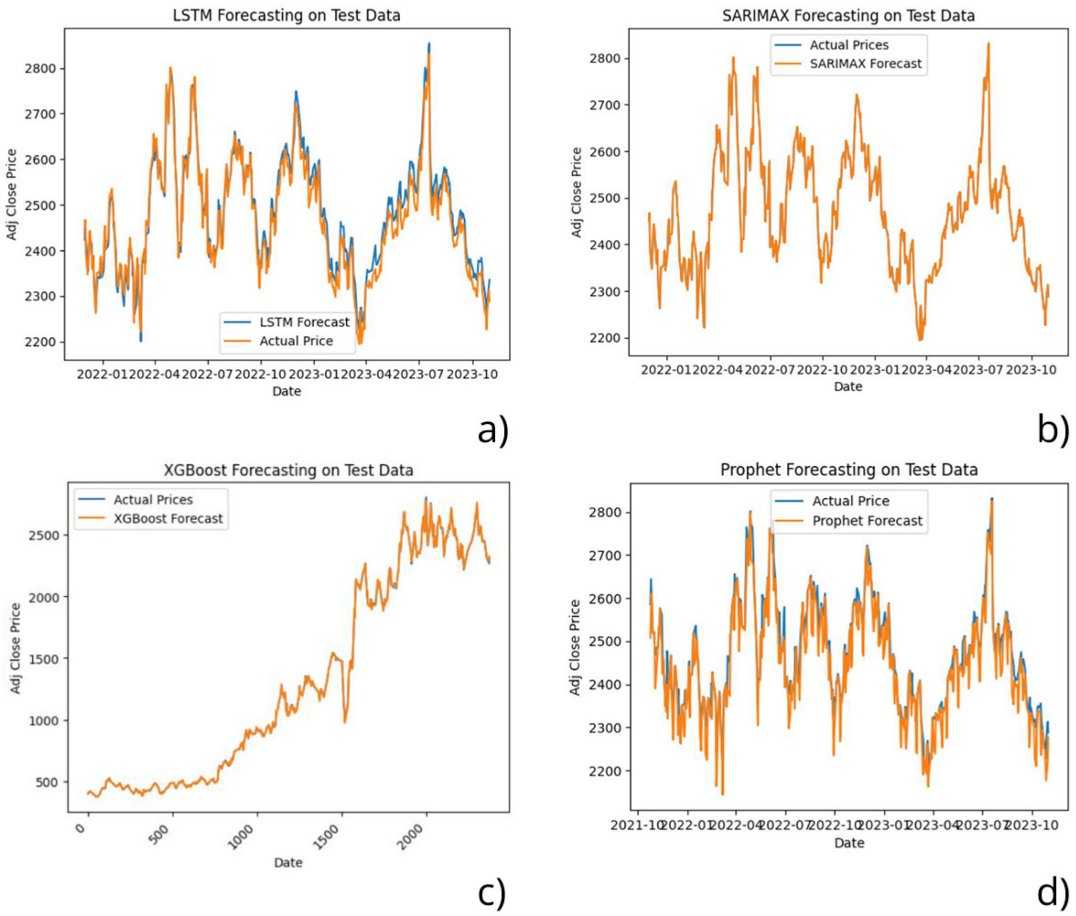

Fig. 5: a) Forecasting plot for LSTM – merged data set b) Forecasting plot of SARIMAX – merged with economical indicators c) Forecasting results of XG Boost d) Forecasting with FB prophet

Source: compiled by the author using Google Colaboratory- FBProphet after integration with economic variables, 2023.; Google Colaboratory- LSTM after integration with economic variables, 2023; Google Colaboratory- SARIMAX after integration with economic variables, 2023; Google Colaboratory- XGBoost after integration with economic variables, 2023.

using a heatmap or correlation matrix. The correlation coefficients indicate the strength of the relationships between the independent variables and the target variable. The correlation coefficient between Gold Price and US dollar to INR conversion is the highest among the economic factors included, at 0.93 and 0.90 respectively. Following that, WPI has a moderate correlation with the ‚Adj Close' at 0.48, while CPI has the lowest correlation coefficient at 0.16. Additionally, the factor Indian 10-year bond price shows an inverse relationship with the target variable, with a correlation coefficient of -0.29.

For LSTM, initial step of importing the necessary libraries was kept similar like previous model. After the pre-processing the data was split again in two ratios to analyse the result efficiently and compare it with previous LSTM model. The model was built using a similar process of developing a sequential neural network. The plot has been depicted in Figure 5a). For SARIMAX, importing the necessary libraries were kept like SARIMAX model before. The model was built from 'pmdarima' to select the best SARIMAX order based on Akaike Information Criterion which was (0,1,2) but again the order provided was (0,0,0) to observe the changes. The seasonality was kept at none. The forecasting has been depicted in Figure 5b). For XGBoost, the initial process of importing the necessary libraries and data pre-processing have been kept like the previous XGBoost model. The train test split has been kept 80 percent of training data and 20 percent of test data. The plotting of the results is depicted in Figure 5c).

Model	Trainnig %	Metric	Before integration with economic variables	After integration with economic variables	Interpretation	Possible Reasons for Change	Considerations for Improvement
LSTM	80%	MSE	569.8537	1366.4344	Increase in MSE indicates reduced prediction accuracy	Overfitting, Complexity from Economic Factors	Further hyperparameter tuning, regularization, feature analysis
		MAE	20.2353	32.2994	Rise in MAE suggests decreased accuracy	Overfitting, Influence of Economic Factors	Model simplification, feature selection, regularization
		R2	0.9639	0.9133	Drop in R2 implies reduced explanatory power	Overfitting, Complexity from Economic Factors	Feature engineering, model architecture adjustments, hyperparameter tuning

Tab. 2 Result analysis of LSTM

Source: compiled by the author using Google Colaboratory- LSTM after integration with economic variables, 2023; Google Colaboratory- LSTM before integration with economic variables, 2023

Model	Metric	Before integration with economic variables	After integration with economic variables	Interpretation	Possible Reasons for Change	Considerations for Improvement
SARIMAX	MSE	691.53	158.38	Significant decrease in MSE	Inclusion of economic factors	N/A (No change needed)
	MAE	25.71	11.26	Significant decrease in MAE	Inclusion of economic factors	N/A (No change needed)
	R2	0.96	0.99	R2 remained perfect after training	Inclusion of economic factors	N/A (No change needed)

Tab. 3 Table 3: Result analysis of SARIMAX

Source: compiled by the author using Google Colaboratory- SARIMAX after integration with economic variables, 2023; Google Colaboratory- SARIMAX before integration with economic variables, 2023

Model	Metric	Before integration with economic variables	After integration with economic variables	Interpretation	Possible Reasons for Change	Considerations for Improvement
FB PROPETH	MSE	2961.07	1912.35	Significant decrease in MSE	Influence of economic factors	Review economic factors, consider model adjustments
	MAE	41.12	25.87	Significant decrease in MAE	Influence of economic factors	Review economic factors, consider model adjustments
	R2	0.81	0.88	Decrease in R2 indicates reduced explanatory power	Influence of economic factors	Review economic factors, consider model adjustments

Tab. 4 Result analysis of Fb prophet

Source: compiled by the author using Google Colaboratory- Fb prophet before integration with economic variables, 2023; Google Colaboratory- FBProphet after integration with economic variables, 2023

Model	Metric	Before integration with economic variables	After integration with economic variables	Interpretation	Possible Reasons for Change	Considerations for Improvement
XGBOOST	MSE	31.35440	15.96573	Decrease in MSE suggests improved prediction accuracy	Influence of exogenous factors	Better preprocessing, inclusion of more factors
	MAE	4.15593	2.54016	Decrease in MAE suggests improved accuracy	Influence of exogenous factors	Better preprocessing, inclusion of more factors
	R2	0.99994	0.99997	Slight increase in R2	Influence of exogenous factors	Better preprocessing, inclusion of more factors

Tab. 5 Result analysis of XGBoost

Source: compiled by the author using Google Colaboratory- XGBoost after integration with economic variables, 2023; Google Colaboratory- XGBoost before integration with economic variables, 2023

For Fb-prophet, the initial steps taken were same as the previous model like importing the libraries and pre-processing approach. plotting of the actual price in comparison with predicted price have been showed in Figure 5d)

After the implementation of phase 1, the ranking of models based on MSE were as follows: LSTM (Rank 2), SARIMAX (Rank 3), FB-Prophet (Rank 4) and XGBoost (Rank 1). Based on R2 Score: LSTM (Rank 2), SARIMAX (Rank 3), FB-Prophet (Rank 4) and XGBoost (Rank 1). After the selection of economic variables and implementation on same models. The ranks of models shifted and is as follows based on MSE: LSTM (Rank 3), SARIMAX (Rank 2), FB-Prophet (Rank 4) and XGBoost (Rank 1). This showed that LSTM was not able to deal with the extra information and the model was overfitting. The analysis of the implementation of LSTM, SARIMAX, FB-prophet and XGBoost has been depicted in the Table 2, 3, 4 and 5 before and after merging the economic factors (gold prices, USD to INR conversion, WPI, CPI and Indian 10-year bond price). The performance of the model has been measured based on mean square error, mean absolute error and R2 score.

4 RESULTS

The LSTM results from Table 2 reveal a notable decline in predictive performance when the model was trained with 80% of the data, despite the inclusion of economic factors. This decline, reflected in the Mean Squared Error (MSE) increasing from 569.8537 to 1366.4344, Mean absolute error increasing from 20.2353 to 32.994 and decrease in R2 score from 0.9636 to 0.9133 prompts a closer examination of the inner workings of the LSTM architecture, particularly focusing on the input gate, forget gate, and output gate.

The input gate, responsible for determining the relevance of new information to be stored in memory, played a crucial role in this scenario. After integration, the LSTM might not have effectively utilized the economic indicators, potentially due to insufficient exposure to diverse instances in the training data. This limited exposure could have hindered the model‘s ability to generalize complex economic patterns, leading to a suboptimal decision-making process by the input gate. Furthermore, the forget gate, responsible for deciding what information from the previous state should be discarded, likely faced challenges in discerning crucial patterns in the absence of a sufficiently rich dataset. Consequently, the LSTM might have unintentionally discarded relevant economic information, resulting in a degradation of predictive accuracy. This shortfall is indicative of the model‘s struggle to capture long-term dependencies in economic trends, a core strength of LSTM architecture. On the other hand, it could also have been due to overfitting of data, as the sample size of data is relatively small and the number of features has increased thus providing excess information which evidently resulted in decline of model's performance.

The performance evaluation of the SARIMAX model in Table 3 indicates a significant improvement when economic indicators are incorporated. Initially trained on 80% of the data without economic factors, the model relied solely on historical patterns for predictions. However, the inclusion of economic indicators such as gold price, Wholesale Price Index (WPI), Consumer Price Index (CPI), US to INR exchange rate, and India‘s 10-year yield bond price resulted in notable enhancements. The decrease in Mean Squared Error (MSE) from

691.53 to 158.38 and Mean Absolute Error (MAE) from 25.71 to 11.26, accompanied by an increase in the R-squared (R2) score from 0.96 to 0.99, underscores the effectiveness of incorporating economic factors. These improvements suggest that the SARIMAX model, characterized by its order parameters (p, d, q), benefitted from the inclusion of economic predictors. The choice of order parameters (0,0,0) signifies no autoregressive, differencing, or moving average components considered in the initial model. While the suggested parameters (p=0, d=1, q=2) based on manual and automatic inspection may lead to a perfect fit with an R2 score of 1, it‘s crucial to consider the implications of overfitting, especially with a small sample size. By opting for (0,0,0) parameters, the SARIMAX model remains parsimonious, avoiding unnecessary complexity that could potentially lead to overfitting. Moreover, it‘s important to note that stock prices are inherently volatile and influenced by numerous unpredictable factors. In such cases, overly complex models may capture noise rather than genuine patterns, leading to poorer generalization performance. The choice of (0,0,0) parameters indicates that SARIMAX primarily accounts for short-term dependencies in the time series data and adjusts to external factors through the inclusion of economic predictors. This approach strikes a balance between model complexity and performance, ensuring robustness in forecasting stock prices.

The results from Table 4 showcase an increase in the predictive performance of Facebook Prophet (FB Prophet) after training with 80% of the data, notably influenced by the inclusion of economic factors. FB Prophet, a renowned time series forecasting model, is adept at handling business problems through the utilization of historical data. Thus, the addition of economic indicators, such as gold price, Wholesale Price Index (WPI), Consumer Price Index (CPI), US to INR exchange rate, and India‘s 10-year yield bond price, led to a decrease in Mean Squared Error (MSE) from 2961.07 to 1912.35 and Mean Absolute Error (MAE) from 41.12 to 25.87. The perfect R-squared (R2) increased to 0.88, indicating a rise in the model‘s explanatory power. FB Prophet operates through the workflow involving the calculation of logistic trend functions, capacity values, yearly seasonality, and the identification of changepoints. The model‘s core equation comprises a trend component (g(t)), seasonality component (s(t)), holiday effect (h(t)), and an error term (εt). It leverages a piecewise function for trend modelling with linear trends and user-specified changepoints. Additionally, FB Prophet utilizes Fourier series for flexible modelling of seasonality patterns. The observed increase in FB Prophet‘s performance suggests that the model, sensitive to the influence of economic factors, may have adapted to the additional complexity introduced by these external variables. To enhance predictive accuracy furthermore, considerations for improvement could be a detailed review of the specific economic factors, potential adjustments to model parameters such as trend components and seasonality, and a strategic handling of anomalies within the economic indicators. This highlights the importance of a nuanced understanding and careful integration of external factors for effective time series forecasting using FB Prophet.

The results of XGBoost from Table 5 depicts that the model demonstrated a significant improvement in predictive accuracy after training with 80% of the data, marked by a decrease in Mean Squared Error (MSE) from 31.3544 to 15.96573, a reduction in Mean Absolute Error (MAE) from 4.15593 to 2.54016, and a slight increase in R-squared (R2) from 0.99994 to 0.99997. This enhancement can be attributed to XGBoost‘s ensemble learning approach, which sequentially trains decision trees and combines their outputs. The model effectively utilized exogenous factors, refining its predictions by compensating for errors made by previous trees and mitigating overfitting through pruning. The mathematical framework, guided by an objective function comprising a Loss function and Regularization, facilitated optimal parameter tuning.

SARIMAX and XGBoost stands out as the most suitable model for forecasting in the provided Indian market data when compared to LSTM and FB Prophet. The models demonstrated a substantial improvement in predictive accuracy, showcasing its adaptability to the influence of exogenous factors. In contrast, LSTM showed sensitivity to dataset size, FB Prophet faced challenges even after the integration of economic factors and remained at R2 score of 0.88. Considering the flexibility, accuracy, and adaptability shown by XGBoost throughout, it emerges as the preferred choice for forecasting in the Indian market.

5 DISCUSSION

The studies Siami-Namini and Namin in 2018 and Fischer and Krauss in 2018 recognise the ability of LSTM to capture long term dependencies in time-series data (Fischer and Krauss, 2018; Siami-Namini and Namin, 2018). These studies accentuate the model's superiority in financial markets. However, performance of LSTM depends heavily on size and quality of training data. During the course of this study, MSE changed from 569.8537 to 1366.4344 highlighting the importance of hyperparameter tuning as emphasised by Gorgolis et. al in 2019 (Gorgolis et al., 2019). The model FB Prophet praised by Taylor and Letham in 2017 for its intuitively adjustable parameters shows a significant rise during this study upon inclusion of economic factors with MSE decreasing from 2961.07 to 1912.35 too (Taylor and Letham, 2017). The decline in performance resonates with findings by Taylor and Letham which indicates model's struggle with the integration of complex exogenous variables and this suggests the need for refined parameter adjustments as highlighted by Hamdani et al. in 2023 (Hamdani et al., 2023).

The study by authors Hyndman and Athanasopoulos in 2018, highlights the robustness of SARIMAX in various scenarios (Hyndman, R.J. and Athanasopoulos, G., 2018). The best order parameters (p=0, d=1, q=2) in this study resonates with recommendations from authors Hyndman and Athanasopoulos but the implementation has been performed on the order (0,0,0). In SARIMAX model the reduction of MSE and MAE is a tantamount to model's better performance. Although SARIMAX is highly effective in incorporating exogenous variables during the course of this study it shows potential for overfitting as indicated by reduction in MSE from 25.71 to 11.26. XGBoost as documented by Chen and Guestrin in 2016, as a powerful tree ensemble model and applied by study from Zhang and Chen in 2021 displays remarkable performance in financial domain (Chen and Guestrin, 2016; Zhang and Chen, 2021). The model's ability to handle exogenous variables is also supported by Zhang et al in 2018 (Zhang et al., 2018). The results of this study echo with Chen and Guestrin's findings as shown by the reduction in MSE from 31.35440 to 15.96573. During this study SARIMAX and XGBOOST emerge as reliable options for forecasting in Indian markets. Considering SARIMAX displays results that are prone to overfitting XGBoost's ability to effectively incorporate exogenous variables and ensemble approach suggests it as preferred model.

6 VALIDITY OF CONCLUSIONS / FUTUREWORK

The study demonstrates that among the selected models, XGBoost and SARIMAX have shown better efficiency and performance for Indian stock market predictions. On the other hand, LSTM and FB-Prophet had higher errors and lower R2 score which generally represents the extent of fitting of data in the model. The models made have been tuned for a single dataset/company in the whole financial market of India which holds different segments of corporations. Thus, it induces a further question like whether these conclusions or the results can be generalized for another company in India or for another country? Or is there any robustness in the models against market anomality's or external shocks?

As mentioned, this study encompasses only one of the components affecting the financial market i.e., economic variables. A study by Sharif et al. in 2015 discusses the firm specific factors like firm size, return on equity and other variables and how it impacts the stock prices in Bahrin (Sharif et al., 2015). Another study by King et al. in 2012 discusses the factors like political events and economic news impacting the oil prices movements (King et al., 2012). In the study by Al Tamimi et al. in 2011, they discussed internal and external factors affecting the stock prices in UAE. Some of these internal factors were performance of company, change in board of directors etc. and external factors included government rules and regulations, market conditions, behaviour of participants and other uncontrollable variables (Al-Tamimi et al., 2011).

In most of the studies, authors have concluded that there is more than one factor involved that is impacting the stock prices in any country. Thus, to generalize the model for any country or for any other segment in India and getting it secure against external shocks, a study will need to incorporate all above-mentioned parameters into one. Even then, further questioning can be done like on what basis will one assign the weights to each factor so that the model can learn without getting biased to any one factor. The models made in this study are based on two factors i.e., model characteristics to deal with error (for instance XGBoost uses previous tree to deal with errors) and other factor impacting Indian financial market. If the future work wants to generalize the results of this study, they will have to take caution as every financial market has their own characteristics, regulations, and impact from economic conditions. For example, share of reliance industries are dependent on economic conditions as it is multi-national company. On the other hand, shares prices of a cosmetic company, food related companies will show different behaviour (for instance, food related firms will be more sensitive towards consumers, thus have a relation with CPI). The development of these models can be generalized only if, the future work includes implementation of these models on a company which is affected by economic indicators with change in data of these variables with each country. Also, the robustness of the model is heavily dependent on accuracy of data involved. Thus, it is a necessity to consider that. The current conclusion of the study is heavily dependent upon the characteristics of models. Thus, if there is external shock involved for instance, a new policy implemented by government that will affect the company's share price, the model will face difficulties in predicting the stock prices as they have done in this study. To make the results more generalized and model more robust further study is needed to be done in future.

The study also has other few shortcomings whilst delivering the subject matter which can be filled in future by another research. Some of these are:

Factors	Rank of countries			
	China	United States	United Kingdom	India
Internet Affordability	34	32	22	28
Internet Quality	10	6	32	16
Electronic Infrastructure	49	5	13	91
Electronic Security	79	43	23	66
Electronic Government	16	2	8	35

Tab. 6 Digital Quality Index 2023

Source: compiled by the author (using Digital Quality Index, 2023)

- The study included basic preprocessing approaches rather than complex processes to get the optimal performance of the desired models.
- The study also interprets the results based on the author's understanding on subject.
- The study focuses on particularly 5 economic indicators from India, some of which are only on monthly basis available like CPI and WPI. This shortcoming can be improved by taking daily prices to improve the model's accuracy and reducing the error further via paid websites or dealers.
- The study also focuses on several of the economic indicators which provides a general understanding of India's economic conditions. The research can be improved further by incorporating more economic indicators and implementing them to the same models.

An additional constraint to the hypothesis lies in the assumption that the limited availability of data and published materials is a primary factor contributing to the scarcity of new hybrid models developed for the Indian market. However, an alternative factor may also be influential, as exemplified by India's placement at the 52nd position among 121 countries in the Digital Quality Index 2023 as assessed by surfshark.com (DQI index, 2023). The evaluation criteria encompass Internet Affordability, Internet Quality, Electronic Infrastructure, Electronic Security, and Electronic Government. In Table 6, the comparison of the top four countries including India in terms of publication is grounded in these criteria. The findings suggest that the paucity of data, particularly regarding the performance of models in the Indian market, may be correlated with India's ranking of 91st out of 121 countries in electronic infrastructure. Despite being the fifth-largest economy and possessing the potential to yield substantial returns for investors, India's digital landscape does not adequately support technological advancement. Furthermore, the decline in Internet affordability from the 8th rank in 2020 to the 28th rank in 2023 raises concerns. This trend implies that even if novel applications or software are developed using new hybrid models, accessibility for users may be constrained. Consequently, this scenario contributes to a diminished interest in the field within the Indian market.

Even after considering these variables, according to annual growth rate of documents, domain of automated trading system is just going to get bigger in future ahead and one of the leading potential economies will be India and other developing countries.

7 CONCLUSION

This research contributes to the understanding of the performance of LSTM, SARIMAX, FB Prophet, and XGBoost models, emphasizing the impact of economic factors on predictive accuracy. The comprehensive exploration into time series forecasting within the dynamic landscape of the Indian financial market has yielded valuable insights, offering practical implications for deploying various algorithms and integrating economic indicators with conventional features for investors, decision makers, traders or academicians. The LSTM model demonstrated a superior performance initially but a notable declined was observed after the integration of economic variables. XGBoost exhibited efficient performance throughout, displaying a decrease in errors, and highlighting its robust ensemble learning capabilities. SARIMAX after incorporating economic indicators, delivered better results with a substantial decrease in both MSE and MAE. FB Prophet also performed relatively well with the new dataset, emphasizing the importance of careful consideration when incorporating economic factors. The findings can be applied in the domain of research, as it provides basic supplementary information on which newer age models for Indian financial market can be developed. The knowledge of the performance of selected models and the visualisation of the established relation between macro-economic variables and time series forecasting models can be used by potential investors, potential traders or decision makers who wants to create a strategy for buying and selling or multiply their investment through algorithmic trading.

Acknowledgements

The publication was supported by the IGA doctoral project (Internal grant agency PEF), namely project IGA-PEF-TP-23-013.

This paper was supported by the project CZ.02.1.01/0.0/0.0/16_017/0002334 Research Infrastructure for Young Scientists, this is co-financed from Operational Programme Research, Development and Education.

REFERENCES

ABU-MOSTAFA, Y. S. and ATIYA, A. F. 1996. Introduction to financial forecasting. *Appl Intell,* 6, 205–213. https://doi.org/10.1007/BF00126626

AGRAWAL, M., KHAN, A. U. and SHUKLA, P. K. 2019. Stock price prediction using technical indicators: a predictive model using optimal deep learning. *Learning,* 6, 7.

AITHAL, P., ACHARYA, D. and GEETHA, M. 2019. Identifying Significant Macroeconomic Indicators for Indian Stock Markets. *IEEE Access*, 7, 143829-143840. https://doi.org/10.1109/ACCESS.2019.2945603

AL-TAMIMI, H. A. H., ALWAN, A. A. and ABDEL RAHMAN, A. A. 2011. Factors Affecting Stock Prices in the UAE Financial Markets. *Journal of Transnational Management,* 16, 3–19. https://doi.org/10.1080/15475778.2011.549441

BATRA, V., KANDPAL, D. and SINHA, R. 2020. Relationship between exchange rate (usd/inr) and stock market indices in India (sensex). *Asian Journal of Research in Banking and Finance,* 10, 1–14.

BHUNIA, A. and DAS, A. 2012. Association between gold prices and stock market returns: Empirical evidence from NSE. *Journal of Exclusive Management Science,* 1, 1–7.

BISWAS, A. 2018. Impact of Reliance Industry Stock Price on NIFTY 50-Granger Causality Test.
RESEARCH REVIEW International Journal of Multidisciplinary, 3(11), 367-376. https://doi.org/10.5281/zenodo.1490556

CHATTERJEE, A., BHOWMICK, H. and SEN, J. 2022. Stock Volatility Prediction using Time Series and Deep Learning Approach. *arXiv*: 2210.02126. https://doi.org/10.48550/arXiv.2210.02126

CHEN, K., ZHOU, Y. and DAI, F. 2015. A LSTM-based method for stock returns prediction: A case study of China stock market. In: *IEEE International Conference on Big Data (Big Data)*. Santa Clara, CA, USA, pp. 2823-2824. https://doi.org/10.1109/BigData.2015.7364089

CHEN, T. and GUESTRIN, C. 2016. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 785–794. https://doi.org/10.1145/2939672.2939785

DASSANAYAKE, W., ARDEKANI, I., JAYAWARDENA, C., SHARIFZADEH, H. and GAMAGE, N. 2021. Forecasting accuracy of Holt-Winters exponential smoothing: Evidence from New Zealand. *New Zealand Journal of Applied Business Research,* 17, 11–30. https://doi.org/10.3316/informit.391329680991168

DHINGRA, K. and KAPIL, S., 2021. Impact of Macroeconomic Variables on Stock Market—An Empirical Study. In: LAKHANPAL, P., MUKHERJEE, J., NAG, B. and TUTEJA, D. (Eds.). *Trade, Investment and Economic Growth: Issues for India and Emerging Economies*. Springer, Singapore, pp. 177–194. https://doi.org/10.1007/978-981-33-6973-3_12

DICKEY, D. A. and FULLER, W. A. 1981. Likelihood Ratio Statistics for Autoregressive Time Series
with a Unit Root. *Econometrica*, 49, 1057–1072. https://doi.org/10.2307/1912517

FARID, S., TASHFEEN, R., MOHSAN, T. and BURHAN, A., 2023. Forecasting stock prices using a data
mining method: Evidence from emerging market. *International Journal of Finance & Economics,* 28, 1911–1917. https://doi.org/10.1002/ijfe.2516

FISCHER, T. and KRAUSS, C. 2018. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research,* 270, 654– 669. https://doi.org/10.1016/j.ejor.2017.11.054

GOOGLE COLABORATORY. 2023. *Fb prophet before integration with economic variables* [online]. https://colab.research.google.com/drive/1MbtF5EsiczjPpseNTRPur0CNRiL3KCwk#scrollTo=6zNcy9mP-djZ [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *FB Prophet after integration with economic variables* [online]. https://colab.research.google.com/drive/1T1Wfgbbdzv1o0xIrIk3Cqd-ZeYvt8rol#scrollTo=CVTm37ixq6MX [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *LSTM after integration with economic variables* [online]. https://colab.research.google.com/drive/10geD107lGWAusxI4VJzA-06ywM_Xe0C4#scrollTo=gtjgxvaXbikH [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *LSTM before integration with economic variables* [online]. https://colab.research.google.com/drive/1KnpEyH-SOnyCGDWH2m9U3oDB4evqcxBm#scrollTo=aZhdNWwNNLCr [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *SARIMAX after integration with economic variables* [online]. https://colab.research.google.com/drive/1ZKlhC1P67NdeyysWn7Ml_rDouC4yvWEB#scrollTo=SYDozCaHrPuK [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *SARIMAX before integration with economic variables* [online]. https://colab.research.google.com/drive/1tBVT3L24QseeGO7VjBK3TuZvK44uCSCJ#scrollTo=ctZInzRBWxcK [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *XGBoost after integration with economic variables* [online]. https://colab.research.google.com/drive/1jadYL8RGEokaIWYo6IhQliYG0R3UR8wc#scrollTo=UJuoBi5FhGFl [accessed 6.10.24].

GOOGLE COLABORATORY. 2023. *XGBoost before integration with economic variables* [online]. https://colab.research.google.com/drive/1jCwnYGKrDhdghVySmDL5mf6gO7odMp#scrollTo=WJe9w8A8lP18 [accessed 6.10.24].

GORGOLIS, N., HATZILYGEROUDIS, I., ISTENES, Z. and GRAD-GYENGE, L. 2019. Hyperparameter Optimization of LSTM Network Models through Genetic Algorithm. In: *10th International Conference on Information, Intelligence, Systems and Applications (IISA)*. Patras, Greece, pp. 1-4, https://doi.org/10.1109/IISA.2019.8900675

GRANGER, C. W. J. 1969. Investigating Causal Relations by Econometric Models and Cross- spectral Methods. *Econometrica,* 37, 424–438. https://doi.org/10.2307/1912791

GUMELAR, A., SETYORINI, H., ADI, D., NILOWARDONO, S., LATIPAH, WIDODO, A., TEGUH WINOWO, A., SULISTYONO, M. and CHRISTINE, E. 2020. Boosting the Accuracy of Stock Market Prediction using XGBoost and Long Short-Term Memory. In: *International Seminar on Application for Technology of Information and Communication (iSemantic)*. Semarang, Indonesia pp. 609–613. https://doi.org/10.1109/iSemantic50169.2020.9234256

HAMDANI, A. F., SWANJAYA, D. and HELILINTAR, R. 2023. Facebook Prophet Model with Bayesian Optimization for USD Index Prediction. *JUITA: Jurnal Informatika*, 11, 293–300. https://doi.org/10.30595/juita.v11i2.17880

HUSSAIN, M., MALIK, A., RASOOL, N., FAYYAZ, M. and MUMTAZ, M. 2012. The Impact of Macroeconomic Variables on Stock Prices: An Empirical Analysis of Karachi Stock Exchange. *Mediterranean Journal of Social Sciences,* 3, 295–312. https://doi.org/10.5901/mjss.2012.v3n3p295

HYNDMAN, R. J. and ATHANASOPOULOS, G. 2018. *Forecasting: principles and practice*. OTexts.

https://books.google.cz/books?hl=en&lr=&id=_bBhDwAAQBAJ&oi=fnd&pg=PA7 &dq=forecasting+principles+and+practice+2018&ots=Tjg-wfWPGK&sig=lFdCll7vOVMtzQ5hShUpgM1KLc0&redir_esc=y#v=onepage&q=sa rima&f=false

JOHANSEN, S. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press. New York.

KARMAKAR, M. 2005. Modeling Conditional Volatility of the Indian Stock Markets. *Vikalpa,* 30, 21–38. https://doi.org/10.1177/0256090920050303

KHASHEI, M. and BIJARI, M. 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing, The Impact of Soft Computing for the Progress of Artificial Intelligence*, 11, 2664–2675. https://doi.org/10.1016/j.asoc.2010.10.015

KING, K., DENG, A. and METZ, D. 2012. An econometric analysis of oil price movements: the role of political events and economic news, financial trading, and market fundamentals. *Bates White Economic Consulting*, 1, 53.

KUMAR, J. P. S., SUNDAR, R. and RAVI, A. 2023. Comparison of stock market prediction performance of ARIMA and RNN-LSTM model - A case study on Indian stock exchange. *Presented at the AIP Conference Proceedings*, 2875, 020010. https://doi.org/10.1063/5.0154124

KUMARIA, A., RAJKAR, A., RAUT, A. and NAIR, R. S. 2023. Forecasting the Indian Financial Markets with LSTM and Price Indicators. In: RAY, K. P., DIXIT, A., ADHIKARI, D. and MATHEW, R. (Eds.). *Proceedings of the 2nd International Conference on Signal and Data Processing*. Springer Nature, Singapore, pp. 395–403. https://doi.org/10.1007/978- 981-99-1410-4_33

LAKSHMANASAMY, T. 2021. The Relationship Between Exchange Rate and Stock Market Volatilities in India: ARCH-GARCH Estimation of the Causal Effects. *International Journal of Finance Research,* 2, 244–259.

LEE, K. J., CHI, A. Y., YOO, S. and JIN, J. J. 2008. Forecasting korean stock price index (kospi) using backpropagationn neural network model bayesian chiao's model and sarima model. *Journal of Management Information and Decision Sciences*, 11.

LITZENBERGER, R., CASTURA, J. and GORELICK, R. 2012. The impacts of automation and high frequency trading on market quality. *Annu. Rev. Financ. Econ.,* 4, 59–98.

LIU, L., PEI, Z., CHEN, P., LUO, H., GAO, Z., FENG, K. and GAN, Z. 2023. An Efficient GAN-Based Multi-classification Approach for Financial Time Series Volatility Trend Prediction. *Int J Comput Intell Syst,* 16, 40. https://doi.org/10.1007/s44196-023-00212-x

LIU, Y., HUANG, S., TIAN, X., ZHANG, F., ZHAO, F. and ZHANG, C. 2024. A stock series prediction model based on variational mode decomposition and dual-channel attention network. *Expert Systems with Applications,* 238, 121708.

MEHTAB, S., SEN, J. and DUTTA, A. 2021. Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. In: THAMPI, S. M., PIRAMUTHU, S., LI, K.-C., BERRETTI, S., WOZNIAK, M. and SINGH, D. (Eds.). *Machine Learning and Metaheuristics Algorithms, and Applications*. Springer, Singapore, pp. 88–106. https://doi.org/10.1007/978-981-16-0419-5_8

NASIRI, H. and EBADZADEH, M. M. 2023. Multi-step-ahead stock price prediction using recurrent fuzzy neural network and variational mode decomposition. *Applied Soft Computing,* 148, 110867.

OBTHONG, M., TANTISANTIWONG, N., JEAMWATTHANACHAI, W. and WILLS, G. 2020. A survey on machine learning for stock price prediction: algorithms and techniques. In: *Presented at the 2nd International Conference on Finance, Economics, Management and IT Business.* (05/05/20 - 06/05/20), pp. 63–71. https://doi.org/10.5220/0009340700630071

PANIGRAHI, A. et al. 2022. Impact of Global and Domestic Economic Variables on 10-Year Indian Government Bond Yield: An Empirical Study. *IUP Journal of Applied Finance; Hyderabad*, 28(2), 5-23.

PATEL, S. A. 2013. Causal Relationship Between Stock Market Indices and Gold Price: Evidence from India. *IUP Journal of Applied Finance,* 19(1), 99-109.

RAHEEM AHMED, R., VVEINHARDT, J., ŠTREIMIKIENĖ, D., GHAURI, S. P. and AHMAD, N. 2017. Estimation of long-run relationship of inflation (CPI & WPI), and oil prices with KSE-100 index: evidence from Johansen multivariate cointegration approach. *Technological and Economic Development of Economy*, 23(4), 567-588. https://doi.org/10.3846/20294913.2017.1289422

SAPRE, A. A. and GORI, S. 2023. The Predicament of Land Acquisition, Displacement and Resettlement: An Analysis of Indian Scenario. *Journal of Asian and African Studies,* 00219096231179651. https://doi.org/10.1177/00219096231179651

SEAH, S. 2022. Untapped Potential in the ASEAN-India Relationship: Climate Change and Green Recovery, In: *ASEAN and India. WORLD SCIENTIFIC*, pp. 227–233. https://doi.org/10.1142/9789811262906_0027

SHARIF, T., PUROHIT, H. and PILLAI, R. 2015. Analysis of Factors Affecting Share Prices: The Case of Bahrain Stock Exchange. *International Journal of Economics and Finance,* 7(3), 207. https://doi.org/10.5539/ijef.v7n3p207

SHARMA, K., BHALLA, R. and GANESAN, G. 2022. *Time Series Forecasting Using FB-Prophet*. Presented at the ACM.

SIAMI-NAMINI, S. and NAMIN, A. S. 2018. Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. *arXiv*: 1803.06386. https://doi.org/10.48550/arXiv.1803.06386

SINGH, P. and BORAH, B. 2014. Forecasting stock index price based on M-factors fuzzy time series and particle swarm optimization. *International Journal of Approximate Reasoning,* 55, 812–833. https://doi.org/10.1016/j.ijar.2013.09.014

Smith, G. n. d. *The Price of Gold and Stock Price Indices for The United States*.

SONKAVDE, G., DHARRAO, D., BONGALE, A., DEOKATE, S., DORESWAMY, D. and BHAT, S. 2023. Forecasting Stock Market Prices Using Machine Learning and Deep Learning Models: A Systematic Review, Performance Analysis and Discussion of Implications. *International Journal of Financial Studies,* 11, 94. https://doi.org/10.3390/ijfs11030094

SRIVASTAVA, S., PANT, M. and GUPTA, V. 2023. Analysis and prediction of Indian stock market: a machine-learning approach. *International Journal of System Assurance Engineering and Management,* 14, 1567–1585.

SUBBURAYAN, B., DHIVYA, N. and ALEX, A. 2021. Empirical Relationship of Macroeconomic Variables and Stock Prices : Indian Stock Market and Japanese Stock Market.

TAYLOR, S. and LETHAM, B. 2017. Forecasting at Scale. *The American Statistician*, 72(1), 37-45. https://doi.org/10.1080/00031305.2017.1380080

ZHANG, D., QIAN, L., MAO, B., HUANG, C. and SI, Y. 2018. A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGBoost. *IEEE Access*, 6, 21020-21031,. https://doi.org/10.1109/ACCESS.2018.2818678

ZHANG, Y. and CHEN, L. 2021. A Study on Forecasting the Default Risk of Bond Based on XGboost Algorithm and Over-Sampling Method. *Theoretical Economics Letters,* 11, 258–267. https://doi.org/10.4236/tel.2021.112019

Reliance Industries Limited (reliance.ns) stock historical prices & data. 2023. *Yahoo! Finance* [online]. https://finance.yahoo.com/quote/RELIANCE.NS/history?p=RELIANCE.NS [Accessed: 01 December 2023].

Digital quality of life index. 2023. *Surfshark.com* [online]. https://surfshark.com/dql2023 [Accessed: 20 December 2023].

Gold futures historical prices. 2023. *Investing.com India* [online]. https://in.investing.com/commodities/gold-historical-data [Accessed: 20 December 2023].

India 10-year Bond Historical Data. 2023. *Investing.com India* [online]. https://in.investing.com/rates-bonds/india-10-year-bond-yield-historical-data [Accessed: 20 December 2023].

India wholesale price index (WPI). 2023. *Investing.com India* [online]. https://in.investing.com/economic-calendar/indian-wpi-inflation-564 [Accessed: 20 December 2023].

India consumer price index (CPI). 2023. *Investing.com India* [online].. Available at: https://in.investing.com/economic-calendar/indian-cpi-973 [Accessed: 20 December 2023].

USD INR historical data. 2023. *Investing.com India* [online]. https://in.investing.com/currencies/usd-inr-historical-data%20 [Accessed: 20 December 2023].

Contact information

Ankit Tripathi: email: xtripat1@.mendelu.cz
Arpit Tripathi: email: xtripat2@.mendelu.cz
Oldrich Trenz: email: oldrich.trenz@mendelu.cz
Pawan Kumar Mishra: email: pawan.mishra@mendelu.cz