

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

TEDIA

International Journal of Innovative Research in SCIENCE | ENGINEERING | TECHNOLOGY

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 6, June 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

## Impact Factor: 8.118

9940 572 462

6381 907 438

 $\bigcirc$ 





| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 6, June 2022

DOI:10.15680/IJIRSET.2022.1106087

## Implementation towards Stock Prediction System used in recent Times

Prof. S. G. Joshi<sup>1</sup>, Ms. Priyanka Dagadkhair<sup>2</sup>

Assistant Professor, Department of Computer Engineering, Vishwa Bharati Academy's College of engineering,

Ahmednagar, India

Research Student, Department of Computer Engineering, Vishwa Bharati Academy's College of engineering,

Ahmednagar, India

**ABSTRACT:** In the financial sector, the stock market and its trends are highly volatile in nature. Recent studieshave shown that news articles and social media analysis can have an immense impact on investors'opinion toward financial markets. Thus, the purpose of this study is to explore the relationship between news sentiment and stock market movement using information from different news agencies, business magazines, and financial portals. This study offers an application of the Bayesian structural time (BST) series model that is more transparent and facilitates better handling of uncertainty than the autoregressive integrated moving average (ARIMA) model and the vector autoregression (VAR) method by using prior information about the structure of the model. One of the main pitfalls of this model is the presumption of linearity. The long short-term memory (LSTM) model is a nonlinear model that can capture various nonlinear structures present in the data set. Proposed methodology is a hybrid model, which combines the LSTM model with the BST model along with the regression component that captures information from different news sources to identify market predictors. The proposed model detects unusual behavior or anomalous pattern of the stock price movement, which makes our model superior compared to the traditional methods.

KEYWORDS: stock market, prediction, SVM, LSTM, Hybrid Approach (Keywords)

#### I. INTRODUCTION

Use Stock Market prediction and analysis is the act of trying to determine the future value of a company stock or other financial instrument traded on annex change. Stock market is the important part of economy of the country and plays a vital role in the growth of the industry and commerce of the country that eventually affects the economy of the country. This system named "Stock Market Prediction" is a web application that aims to predict stock market value using SVM, LSTM, Linear Regression. This project is intended to solve the economic dilemma created in individuals that wants to invest in Stock Market. With the growth of the Internet, social networks, and online social interactions, getting daily user predictions to make a stronger model that will benefit everyone In earlier 2009, Stock market forecasting was done by ARIMA Model. ARIMA, short for 'Auto Regressive Integrated Moving Average' . It is Time series Model where forecast is done by time series using the series past values. The Model is designed for statistical problems. Though model is adept at modelling seasonality and trends, outliers are difficult to forecast for ARIMA for the very reason that they lie outside of the general trend as captured by the model. The sentiment analysis cannot be done with ARIMA model . The system will be implemented to design and develop a system to detect sentiment analysis for stock investment through news article and social media.

#### **II. LITERATURE REVIEW**

1. Ray, Paramita, Bhaswati Ganguli, and Amlan Chakrabarti. "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price."

Authors propose a hybrid model, which combines the LSTM model with the BST model along with the regression component that captures information from different news sources to identify market predictors. The proposed model detects unusual behavior or anomalous pattern of the stock price movement, which makes our model superior compared to the traditional methods. Our new hybrid model accumulates error with lower rates and shows a remarkable performance over some of the other existing hybrid models, such as AR-MLP, ARIMA-LSTM, and VAR-LSTM model.[1]



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 6, June 2022

DOI:10.15680/IJIRSET.2022.1106087

#### 2. Ching-Te Wang, "The Prediction System for Data Analysis of Stock Market by Using Genetic Algorithm"

In this paper, a web robot scheme to crawl the data and collect the information, which is provided for the analysis procedure. The system can predict the stock value for the users. After set the suitable rule, the web robot can collect data as soon as possible. To speed the computation between data crawling and data analysis, the system designs the structure of centralized distributive web robot system, which distinguishes web robots, data server and analysis server. Therefore, the system makes the process automatically and enhances the reaction rate efficiently. The main characteristic of this system is to collect data automatically by using web robots and 2 establish regular expressions, X path to analyzing web pages. From the experimental results, the system has shown a better performance. Consequently, the method can crawl the valuable data, analyze huge information efficiently and provide the function to predict the prices in the stock market[2].

#### 3. Sahaj Singh Maini, "Stock Market Prediction using Data Mining Techniques":

In this paper, the performed predictive analysis on Dow Jones Industrial Average Index, which consists of thirty companies and is majorly owned by S P Global. This is crucial to the traders as such analysis can influence the decision making with regard to buying or selling an instrument in a positive manner. We have discussed two statistical machine learning models, namely Random Forest Model and Support Vector Machine which are used to provide a reliable prediction of stock market trends based on historical data. On the basis of the results obtained; we can say that both the models exhibited notable performance in predicting the direction of the stock index. The Random Forest model using a 1-gram model for text analysis produced an accuracy of 84.33 [3].

### 4. Chang SimVui, Gan Kim Soon, Chin Kim On, "A Review of Stock Market Prediction with Artificial Neural Network (ANN)":

This paper provides a review on the application of ANN in stock market prediction. Although NN has shown acceptable results, many researchers are still trying to improve the accuracy of the stock market prediction by using a hybrid method, and by taking into account more external factors in order to generate more accurate prediction. It is a known fact that the dynamic stock market world is non-linear, volatile and subject to influences by so many external factors. This paper can be used as an introductory material to those who are interested to work on stock market prediction using NN [4].

#### 5. RaduIacomin, "Stock Market Prediction":

This study generated a new algorithm on predicting the stock markets. As it was shown in this paper, there are several methods on predicting signals through machine learning algorithms and numerical methods. PCASVM was implemented to both eliminate the false predictions and to determine what features are important. Comparing to the simple methods from SVM and volving to GASVM and PCASVM, the solution to the main problem and sub-issues is more efficient and shows promising results for a real prediction using recent data sets[5].

#### 6. A Study On Various Data Mining Techniques For Crop Yield Prediction

India is a country where agriculture and agriculture related industries are the major source of living for the people. Agriculture is a major source of economy of the country. It is also one of the country which suer from major natural calamities like drought or good which damages the crop. As Prediction of crop deals with large set of database thus making this prediction system a perfect candidate for application of data mining. Through data mining we extract the knowledge from the huge size of data. This paper presents the study about the various data mining techniques used for predicting the crop yield. The success of any crop yield prediction system heavily relies on how accurately the features have been extracted and how appropriately classifiers have been employed. This paper summarizes the results obtained by various algorithms which are being used by various authors for crop yield prediction, with their accuracy and recommendation.[6]

#### 7. System For Agriculture Recommendation Using Data Mining

Agriculture in today's life is not like as our forefather done. The strong Climatic changes due to many reasons like global warming cause difficulty to understand climatic conditions. So the farmers unable to understand which crop to select by which the production will improve. By understanding soil and climate conditions by using these data mining system farmers will be able to take right crop at right place which will improve yields. So it is easy for farmers to decide which crop to take in unpredictable climate conditions. This project will help to solve these agriculture problems using data mining algorithms. Algorithms like Genetic Algorithm (GA), Association Rule Mapping (ARM). Knowledge Discovery in Databases (KDD) is the centroid of this project from which we can really understand the concept of data mining.[7] etc.



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 6, June 2022

DOI:10.15680/IJIRSET.2022.1106087

#### **III. PROPOSED METHOLOGY**

The system is designed to predict the stock values .The outline for system can be jotted down as :

1. Time series forecasting consists of a research area designed to solve various problems, mainly in the financial area.

2. nonlinear regression problems by constructing the input-output mapping function.

3. The least squares support vector regression (LSSVR) algorithm is a further development of SVR and its use considerably reduces computational complexity and increases efficiency compared to standard SVR.

4. The Firefly Algorithm (FA), which is a nature-inspired metaheuristic method, has recently per formed extremely well in solving various optimization problems.

#### A. Prediction system for Machine Learning

Machine Learning (ML) is that field of computer science with the help of which computer systems can provide sense to data in much the same way as human beings do. In simple words, ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method. The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention Predicting how the stock market will perform is one of the most difficult things to do. There are so many factors involved in the prediction - physical factors vs. physiological, rational and irrational behavior, etc. All these aspects combine to make share prices volatile and very difficult to predict with a high degree of accuracy.



#### Fig 1: Prediction System

B. System Architecture for system

- Data Sources
- This project attempts to predict the stock value with respect to the stock's previous value and trends. It requires historic data of stock market as the project also emphasizes on data mining techniques. So, it is necessary to have a trusted source having relevant and necessary data required for the prediction
- Dataset Creation
- First of all, a dataset is created for training the Artificial neural network. In this study, daily stock
- market data Open (10 am), High (1 pm), and Closing value (3 pm) have been collected for the duration of four years (January 1, 2015 to March 29, 2019) from the NSE of India[3]. The collected data are arranged according to the format for the library we use for training. The dataset should be of exact format that FANN specifies. It includes number of training pair, number of input and number of outputs in the first line of the data set file and data from the second line.

L SET

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 6, June 2022

DOI:10.15680/IJIRSET.2022.1106087



Architecture

- Data Normalization
- The data is normalized before being input. The input vectors of the training data are normalized such that all the features are zero-mean and unit variance. The target values are normalized using minmax function such that all the values are converted into the values within the range of 0 to 1. The minimum value is represented by 0 and the maximum value is represented by 1.

#### V. STOCK PREDICTION SYSTEM

Hybrid approach to use LSTM and Bayesian for Prediction The system applies the BST timeseries model with feature selection for forecasting. Bayesian structural time series (BSTS) model is a statistical technique used for feature selection, time series forecasting, nowcasting, inferring causal impact and other applications. The model is designed to work with time series data. The model has also promising application in the field of analytical marketing. The system also uses LSTM for more accuracy from data from social media .LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs.



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

||Volume 11, Issue 6, June 2022||

DOI:10.15680/IJIRSET.2022.1106087

VI. RESULTS AND DISCUSSIONS

📲 PRC 🕘 Rep 🛑 Buil 📴 Selc 🞯 how 🛶 s 🗙 👷 Int	lo 🕘 Nev 🙆 te	ni 🕲 inst 🛛	🛭 Hor 😂 Hor	) hov (@ 1	Rep   @ Sto   @ Mic	@ 100 g	🕲 Whi 🙆 dati	± Dov  ⊗	0   +	
$\leftrightarrow \rightarrow \mathbf{C} \ \mathbf{\hat{\Omega}} \ \mathbf{\hat{O}}$ localhost:8502								± ± •		
Sto	ck Pric	e Pre	dicti	on						Ξ
EnterCom	Jany Name									
AXISBA	NK.NS									
Data f	from 2011	to 2022								
	High	Low	Open	Close	Volume	Adj C				
count	2,788.0000	2,788.0000	2,788.0000	2,788.0000	2,788.0000	2,788.0				
mean	482.0343	468.4949	475.6134	475.1773	11,348,652.9921	468.3				
std	187.1424	183.2140	185.3756	185.2111	9,701,729.7369	189.8				
min	163,0000	152.6800	156.5600	156.5700	0.0000	148.4				
25%	279.1950	272.2575	274.8950	275.7550	5,978,261.5000	260.1				
50%	503.5250	490.0500	496.9000	496.6000	8,551,323.5000	489.0				
75%	621.8250	603,5375	613.1750	612.8375	12,812,866.2500	611.1				
ITUAX	865.9000	825.4500	848.4000	845.1000	120,541,914.0000	845.1				
AXI 800- 790-	SBANK	(.NS o	comp	any cl	nart Ma M	W				

Volume	Adj Close**	Close*	Low	High	Dpen	Data
6,019,100	18.70	18.70	18.39	18.83	18.82	May 20, 2022
9,392,800	18.63	18.63	18.22	18.83	18.36	May 19, 2022
10,672,000	18.78	18.78	18.76	19.39	19.32	May 18, 2022
7,184,600	19.61	19.61	19.40	19.62	19.42	May 17, 2022
11,057,600	18.94	18.94	18.90	19.18	19.13	May 16, 2022
14,033,200	19.45	19.45	19.19	19.51	19.30	May 13, 2022
13,434,100	19.53	19.53	19.12	19.65	19.16	May 12, 2022
8,503,200	19.41	19.41	19.38	19.77	19.59	May 11, 2022
18,342,600	19.76	19.76	19.52	20.12	20.02	May 10, 2022
14,538,200	19.82	19.82	19.76	20.14	20.04	May 09, 2022
6,614,900	19.85	19.85	19.71	20.00	19.96	May 06, 2022
9,577,100	20.31	20.31	20.11	20.79	20.60	May 05, 2022
7.519,400	20.61	20.61	20.04	20.66	20.10	May 04, 2022
7,716,800	20.22	20.22	20.13	20.29	20.19	May 03, 2022
15,073,900	20.15	20.15	19.84	20.22	19.88	May 02, 2022
12,324,800	19.87	19.87	19.87	20.30	20.24	Apr 29, 2022
24,229,100	20.50	20.50	20.35	20.66	20.41	Apr 28, 2022
9,754,100	20.12	20.12	19.99	20.32	19.99	Apr 27, 2022
13,027,100	20.12	20.12	20.06	20.38	20.25	Apr 26, 2022

Fig3 : Axis Bank Dataset



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

||Volume 11, Issue 6, June 2022||

DOI:10.15680/IJIRSET.2022.1106087



Fig.4. Axis Bank 100 Days average Price



Fig.5. Axis Bank 100 Days vs 200 days average Price



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

||Volume 11, Issue 6, June 2022||

| DOI:10.15680/IJIRSET.2022.1106087|



Fig. 6 Axis Bank Predicted Price

an end the same time such the same time			the late and la	-	1.	Bar Life Mar Life Ann		
and the state of t						ue l'es sus l'es air-	6.103	
								0 x · · · U · · ·
								=
	Sto.	ale Dei	DE DE	dicti				
	Sto	CK Pri	ce Pre	alcui	оп			
	Enter Com	sany Name						
	INFYS							
	Data	rom 2011	to 2022					
		High	Low	Open	Close	Volume	Adi C	
		2,957.0000	2,867.0000	2,857.0000	2,867.0000	2,867.0000	2,867.0	
	mean	9.9673	9.8180	9.9032	9,9044	11,656,102.2323	8.1	
	std	4.7749	4.6978	4.7369	4.7376	7,888,454.9235	5.0	
		4.8263	4,7413	4.7850	4.7712	1,068,400.0000		
		7.2506	7.1150	7_1775	7.1925	7,195,600.0000	5.6	
		8.4838	8.3200	8.4100	8.3988	9,813,600.0000	7.0	
		10.3725	10.2075	10.3000	10.2925	13,914,650.0000	9.6	
		26.3900	25,5800	26.1500	26.2000	147,591,200.0000	26.3	
	INF	Y com	pany	chart				
		the second states						
	25						1.	
						N	M I	
						1.		



Determinent     Construction     Construction     Add C       mer     2487.000     248.000	Data	from 2011	to 2022	-					
High     Low     Open     Class     Value     All       000     2007 000     2007 000     2007 000     2007 000     2007 000       001     4007 000     2007 000     2007 000     2007 000     2007 000       001     41706     4007 100     2007 000     2007 000     2017       001     41706     41708     41708     41708     41708     41708       205     21206     7120     7175     717	Data	10111 2011	10 2022						=
2471000     24870000     24870000     2487000     24871000     24871000     24871000     24871000     24871     1145412323     51       01     4.7789     4.5789     4.7789     4.7789     5.8717     5.871     5.871     5.871     5.971									
1017     1019     9.002     9.004     11.064.04.023     1       101     4.1766     4.1706     4.170     1.064.04.023     1       101     4.1766     4.1706     4.1712     1.084.04.023     1       1205     7.1713     7.1713     7.1713     7.1713     7.1713     7.1713       1706     1.1173     1.03.000     1.03.000     1.03.000     1.03.000     1.03.000       1706     1.1173     1.03.000     3.13.14.060.0000     1.4     1.03.000		2,867.0000	2,867.0000	2,867.0000	2,867.0000	2,867.0000	2,867.0		
101 4.7700 4.8670 4.770 7.884.646235 5.6   101 4.770 7.840 4.770 7.840.646235 5.6   101 7.250 7.1130 7.277 7.1822 7.840.660300 5.7   101 7.250 7.1130 7.277 7.1822 7.840.660300 5.7   101 1.8775 1.82075 1.834.660300 5.4   101 1.8775 1.82075 1.834.660300 5.4   102 2.8380 2.5380 26.500 1.97.205 1.97.206   INFFY company chart		9.9873	9.8180	9.9032	9.9044	11,656,102.2323			
min     4.1215     4.1313     4.1310     4.1712     1.134000000     3.1       20%     2.1206     7.1215     7.1275		4.7749	4,6978	4.7369	4,7376	7,888,454.9235			
200 1200 1307 13175		4.8263	4.7413	4.7850		1,068,400.0000			
NPh     44439     8.1200     4.4100     X11304     X112000000     T/       max     24.300     32.500     32.600     32.400     32.400     32.400     32.400       INFY company chart		7.2506	7.1150		7.1925	7,195,600.0000			
7%     18.275     10.2675     10.2625     10.2455     0.2456000     94       max     24.3000     25.5000     26.2000     147.291.201.00000     26.2       INFY company chart		8.4838	8.3200	8.4100	5.3355	9,813,600.0000			
Name     24,100     25,2000     247,391,200,000     242       INFY company chart     No     No<		10.3725	10.2075	10.3000	10.2925	13,914,650.0000			
INFY company chart									
- -	INF	26.3900 Y com	25.5800 pany	26.1500 chart	26.2000	147,591,200.0000	26.1		
5 JW 1	max INF 25-	26.3900 Y com	pany	chart	26.2000	147,591,200.0000	26.5		
25- -	75- 25-	26.3900	pany	chart	26.2000	147,591,200.0000	26.2		
23-	175 K	26.3900	pany	chart	26.2000	147,591,200.0000	262		
	75 - 29 -	26.3900	pany	chart	26.2000	147,591,200.0000	265		
	75 - 28 - 23 -	26.3900	pany	chart	26.2000	147,591,000,0000	257		
and the second s	75 - 26 - 38 -	26.3900	pany	chart	25.2000	147,591,000,0000	265		
" we want many	m314 1755 - 25 - 25 -	26.3900 TY com	25.500 pany	chart	25,2000	147,093,000,0000	265		

Fig 8: Infosys Price Char



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 6, June 2022

DOI:10.15680/IJIRSET.2022.1106087



Fig 8: Infosys



Fig 8:Infosys



Fig 8: Infosys Predicted price

#### VI CONCLUSION

The successful prediction of a stock's future price could yield significant profit. The efficient market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable. The plunging stock market has intensified the debate over the future of the bull market, and these three gurus have widely differing views. In this project, demonstrating a



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

#### Volume 11, Issue 6, June 2022

#### DOI:10.15680/IJIRSET.2022.1106087

machine learning approach to predict stock market trend using different neural networks. Result will show how it can use history data to predict stock movement with reasonable accuracy. Also, with T test result analysis we will conclude that LSTM performs better compare to Back propagation and SVM. For this implementation, we would like to conclude that if we incorporate all the factors that affect stock performance and feed them to neural network with proper data preprocessing and filtering, after training the network we will be able to have a model which can predict stock momentum very accurately and this can result into better stock forecasting and profit for financial firms. selfcontained.

#### REFERENCES

[1] Ray, Paramita, Bhaswati Ganguli, and Amlan Chakrabarti. "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price." IEEE Transactions on Computational Social Systems (2021).

[2] Ching-Te Wang," The Prediction System for Data Analysis of Stock Market by Using Genetic Algorithm" pp. 1721-1725, 2015 IEEE.

[3] Sahaj Singh Maini, "Stock Market Prediction using Data Mining Techniques" Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017) pp. 654-661, 2017 IEEE.

[4] Chang SimVui, Gan Kim Soon, Chin Kim On, "A Review of Stock Market Prediction with Artifi cial Neural Network (ANN)" 2013 IEEE International Conference on Control System, Computing and Engineering, 29Nov.-1Dec. 2013, Penang, Malaysia, pp.477-482.

[5] RaduIacomin, "Stock Market Prediction" 2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania.

[6] Yogesh Gandge, Sandhya, "A Study Of Various Data Mining Techniques For Crop Yield Prediction", IEEE, 2017.[7] Omkar Chikane, Nilesh Dumbre, Gitesh More, "System For Agriculture Recommendation Using

Data Mining", IERJ, Dec 2015





Impact Factor: 8.118





## INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com