IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 27, 2019, accepted September 26, 2019, date of publication October 4, 2019, date of current version October 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2945603

Identifying Significant Macroeconomic Indicators for Indian Stock Markets

PRAKASH K. AITHAL, (Member, IEEE), DINESH ACHARYA U., AND GEETHA M.
Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India
Corresponding author: Geetha M. (geetha.maiya@manipal.edu)

ABSTRACT The macroeconomic indicators play a major role in all the stock markets, and they vary from nation to nation. This paper identifies the influence of macroeconomic indicators on the National Stock Exchange (NSE) and the Bombay Stock Exchange (BSE) of India. Total of forty-four macroeconomic indicators for eight years from the year 2011 to 2018 are considered in this study. The macroeconomic factors are aggregated and considered in average monthly form. The proposed method finds the correlation matrix of all considered macroeconomic indicators. The need for dimensionality reduction and the existence of multicollinearity are proven using validation techniques such as the Kaiser-Meyer-Olkin and Bartlett tests. The Principal Component Analysis (PCA) method is used to reduce the dimensionality to seven factors and then PCA with the varimax rotation method is applied to find factors with maximum variation. In addition, the influence of these seven factors on the NSE Nifty and BSE SENSEX indices are analyzed using regression. Finally, an Artificial Neural Network is used to predict stock market movement with the help of macroeconomic indicators. Accuracy of 92% and 87% are obtained on NSE NIFTY and BSE SENSEX respectively.

INDEX TERMS Decision support systems, knowledge discovery, macroeconomic indicators, principal component analysis, artificial neural network, data mining.

I. INTRODUCTION

There are many fundamental theories in finance. For example, the Markowitz risk-return theory, Sharpe's Capital Asset Pricing Model(CAPM), and Arbitrage Price Theory(APT) are prominent among them. The Markowitz model proposes that one must build a portfolio i.e., a collection of stocks, rather than investing in individual stocks. The portfolio is collection of different stocks. The portfolio risk will always be less than the individual stocks forming the portfolio, and return will be an average of returns of the constituting stocks of the portfolio. The CAPM model is an extension of the Markowitz risk-return theory. In the CAPM market risk, or systematic risk is the only risk that affects the stock market, and company-specific risk is reduced to zero by diversification. There are two types in the APT model they are single factor and multi-factor APT. Unlike the CAPM, in a single factor APT, the APT user can select the factor. In multi-factor APT, more than one macroeconomic or financial indicators are considered for study. The Fama French model is a three-factor model. The general APT model does not require returns to be normally distributed, and investors need not be mean-variance optimizers. The APT proposes that the stock market is dependent on macroeconomic indicators.

The influence of macroeconomic indicators on the stock market is inevitable. However, macroeconomic factors affecting one nation need not affect the another nation because different macroeconomic indicators capture different information, some macroeconomic indicators are global in scope, and some are specific to a given nation, and others are unique to specific industries. If a nation is not connected economically with the world, then the influence of macroeconomic indicators with the global scope is either minimal or nonexistent. The proposed paper utilizes the following techniques

- Knowledge Discovery from Data
- Pre-Dimensionality Reduction methods such as
  - Correlation Matrix
  - Bartlett Test
  - Kaiser Meyer Olkin(KMO) Test
- Dimensionality Reduction methods such as
  - PCA
  - K1-Kaiser Method
  - Scree Test
  - PCA with varimax rotation

The associate editor coordinating the review of this manuscript and approving it for publication was Xiping Hu.

- Unit Root Test and measure such as
  - Augmented Dickey-Fuller(ADF) Test
  - Vector Error Correction Model(VECM)
- Data Mining techniques such as
  - Regression Analysis
  - ANN

The above techniques are explained in the following subsection.

A. KNOWLEDGE DISCOVERY FROM DATA

Knowledge Discovery from Data(KDD) as shown in Figure.1. includes data collection, data cleaning, data transformation, data mining, result evaluation, result analysis, and result presentation. Data Mining is a tool to analyze data to extract nontrivial information. The data are collected from the different sources. Data cleaning steps include missing value imputation and outlier removal, and data transformation involves converting data to make it suitable for a given data mining task.

B. PRE-DIMENSIONALITY REDUCTION

The need for dimensionality reduction is identified by applying the correlation matrix, the Bartlett test, and the KMO test.

1) CORRELATION MATRIX

Correlation matrix is obtained by computing the Pearson correlation coefficient formula given in Equation-1.

$$\rho_{xy} = \frac{S_{xy}}{S_x * S_y} \tag{1}$$

Here S_{xy} is the sample covariance of x and y and $S_x and S_y$ are sample standard deviations of variables x and y respectively.

The correlation coefficient value varies between -1 to 1. The value -1 signifies that if the value of X increases, then the value of Y decreases which is a perfect negative correlation. In contrast, a value of 1 signifies if the value of X increases, then the value of Y also increases which is a perfect positive correlation. 0 signifies there is no correlation. Any intermediate value between -1 and 1 indicates that the X and Y variables vary proportionately. If there are many highly positive or highly negative values are present in the correlation matrix, then there are chances of the existence of multicollinearity. Multicollinearity is a situation in which multiple variables are interdependent. Bartlett and KMO tests are used to identify multicollinearity. If multicollinearity is present in the system of variables, then dimensionality reduction technique such as PCA can be applied.

2) BARTLETT TEST

The Bartlett test is used to determine if 'k' variables have the same variance. Equation-2 depicts the Bartlett formula.

$$B = \frac{(N-k)\ln(\frac{\sum_{i=1}^{k}(n_i-1)*s_i^2}{(N-k)}) - \sum_{i=1}^{k}(n_i-1)*\ln(s_i^2)}{1+\frac{1}{3*(k-1)}*[(\sum_{i=1}^{k}\frac{1}{(n_i-1)}) - \frac{1}{(N-k)}]} \tag{2}$$

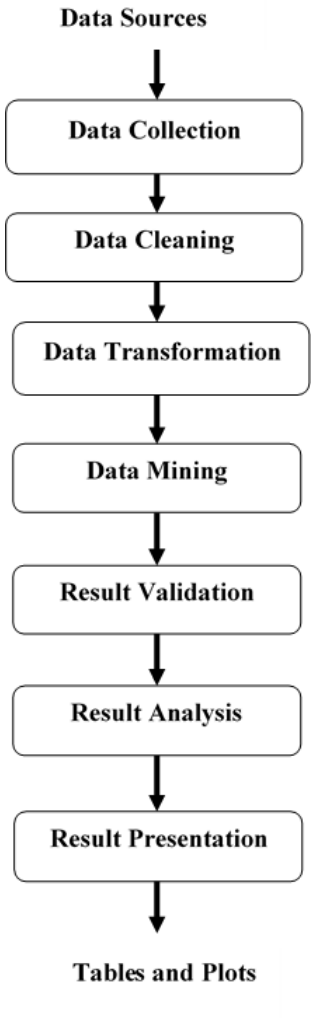

FIGURE 1. KDD steps.

Here N is the sum of all sample sizes, k is the number of normally distributed populations, s_i^2 is the variance, and n_i the number of sample items.

Note that the factor analysis should not be applied if Bartlett test is proved insignificant. Bartlett test utilizes a chi-square test, which is well suited for normal distributions or a large number of variables.

3) KMO TEST

The KMO test validates the adequacy of the sample. Equation-3 depicts the KMO formula.

$$KMO_j = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} u_{ij}} \tag{3}$$

Here r_{ij} is the correlation matrix and u_{ij} is partial covariance matrix.

Researchers have different opinions about the significant value of KMO. One school of thought proposes that the KMO value should be greater than 0.5 while others take the opinion that the KMO value should be greater than 0.6 or 0.7.

The KMO test cannot be applied if there are variables that can deduce other variables perfectly. In other words, KMO test cannot be applied if two variables are perfectly correlated, which occurs when the correlation matrix is nonpositive definite.

C. DIMENSIONALITY REDUCTION

The dimensionality is reduced using the PCA algorithm and related macroeconomic indicators are grouped together into factors. The number of factors to be retained are determined through K1-Kaiser and Scree tests. The prominence of each macroeconomic indicator within a factor is identified using PCA with verimax.

1) PCA

PCA, which is an unsupervised factor analysis method, is a dimensionality reduction technique that takes an input correlation matrix and produces eigenvalue and eigenvector. There are two types of factor analysis i.e. confirmatory and exploratory factor analysis. In confirmatory factor analysis number of factors is known. In contrast, the number of factors is unknown in exploratory factor analysis. In other words, confirmatory factor analysis is supervised factor analysis, and exploratory factor analysis is unsupervised factor analysis. A detailed PCA tutorial is given in literature [1].

2) K1-KAISER

The K1-Kaiser method is used to retain the factors. According to K1-Kaiser method, one must retain the factors that have an eigenvalue greater than one. It is the simplest of the methods to retain the factors.

3) SCREE TEST

The scree test is another method used to retain the factors. According to this test, one must retain factors until there is a long tail in the graphical scree plot. Note that there will be breakpoints in the scree plot until there is a long tail, and these breakpoints distinguish between the stronger factors and less relevant factors.

4) PCA WITH VARIMAX ROTATION

PCA cannot tackle the multicollinearity problem, Thus, PCA with varimax is utilized. In varimax, the rotation axis is rotated orthogonally to solve the multicollinearity problem.

D. UNIT ROOT TEST AND VECM

The unit root test is performed to determine which method to apply to the data. The ADF test is a widely used representative unit root test.

1) ADF TEST

The ADF test finds the autocorrelation of a variable. If the p-value of the ADF test is significant then we say that the variable under study is stationary else non-stationary. If the dependent and independent variables are stationary then simple regression can be performed. If the dependent and independent variables are stationary at first difference and residual is stationary at an initial level, then cointegration between variables exists. If there is only one endogenous variable, an error correction model should be utilized; otherwise, a VECM should be implemented.

2) VECM

VECM is applied when multiple endogenous variables are nonstationary at an initial level when the residual is stationary. They estimate the short and long term effects of one time series over the another.

E. DATA MINING

Data mining is an important step in KDD. Several data mining techniques are available for classification, association, and clustering. The regression analysis and ANNs are considered classification techniques.

1) REGRESSION ANALYSIS

Regression determines the relationship between a dependent variable and independent variable under the assumption that extraneous variables do not affect the equation.

2) ARTIFICIAL NEURAL NETWORK

An ANN is a model that mimics the human brain. The ANN consists of an input layer, hidden layers, and an output layer. The input layer passes the input values to the hidden layers, and the hidden layers pass the result to the output layer. All neurons in an ANN are connected by at least one connection. Neurons are activated by an activation function, which can be a binary step, linear, or nonlinear function.

Note that the binary step function can take only two values: yes or no. A linear function makes the ANN equivalent to a linear regression model: However, a linear function does not allow backpropagation.

The nonlinear activation functions include the sigmoid, hyperbolic tangent, Rectified Linear Unit(ReLU), softmax, and swish functions. The sigmoid function is heavy, while the ReLU function is efficient. The sigmoid function allows all types of integers, while the basic ReLU function fails for negative integers: However, there are ReLU variants that can work with negative values.

Rest of the paper is organized as follows A literature review is provided in, Section II, the methodology is described in Section III, The results are discussed in Section IV, and Section V concludes the paper and gives the future direction.

II. LITERATURE REVIEW

Previous studies have investigated the effect of macroeconomic indicators on the performance of the index. These studies considered the stock market of various nations and a variety of macroeconomic indicators.

For example, Flavin and Wickens [2] studied the influence of macroeconomic factor retail price index on the stock market in the United Kingdom. He observed that Henry

Markowitz's mean-variance analysis model and Sharpe's capital asset pricing model assume a constant portfolio frontier. In reality, the portfolio frontier varies with time due to the effect of macroeconomic factors on stock market performance. The asset pricing theory assumes that stock market performance is influenced by multiple factors referred to as macroeconomic and microeconomic indicators.

Zarina [3] studied the influence of Unexpected Gross Domestic Product(UGDP), Unexpected Total Trade (UTT), and Unexpected Market Returns UMR) relative to the stock markets in Malaysia, Singapore, Thailand, Indonesia, and the Philippines. She found that the UGDP, UTT, and UMR have very little influence on these stock markets.

Xue CHEN and Xuejun JIN [4] studied the Chinese stock market and following macroeconomic indicators risk premium, credit supply, demand factor, economic policy measures and they found that the considered macroeconomic factors influence the Chinese stock market. Note that the study undertaken is not on the pure macroeconomic indicators.

Sudhakaran and Balasubramanian [5] considered the macroeconomic indicators inflation rate, Foreign Direct Investment (FDI), money supply, Foreign Exchange(Forex), industrial production index, and foreign portfolio investment. They selected these indicators because it has not been considered in the previous studies. They identified that FDI and Forex influence the Indian stock market. The macroeconomic indicators were selected randomly.

Claessens and Kose [6] surveyed the influence of macroeconomic indicators exchange rates and interest rates on the international stock markets. They concluded that some researchers think that, due to limited knowledge of macroeconomic indicators the stock markets vary randomly while some others argue that humans have thorough understanding of the influences of macroeconomic indicators on stock markets.

Zhang *et al.* [7] considered the macroeconomic indicators Gross Domestic Product(GDP), inflation, ten-year bond yield, credit spread, and short term interest rate and studied their influence on stock markets in United States of America and United Kingdom. They concluded that macroeconomic factors influence these stock markets.

Gurloveleen and Bhatia [8] have undertaken India BSE 500 stock index for 150 manufacturing firms. They studied the effect of ten macroeconomic indicators namely, crude oil price, call money rate, exchange rate, Forex, broad money, foreign institutional investors, gross fiscal deficit, industrial production index, and inflation rate on the manufacturing firms stock index BSE 500 and conclude that the stock market is weak-form efficient and macroeconomic indicators do not influence the stock market.

Parmar [9] explored the possible influence of seven macroeconomic indicators namely, reserve rate,repo rate, industrial production index, exchange rate, consumer price index, gold rate, and oil rate on the BSE SENSEX the Indian stock index. She concluded that, local macroeconomic indicators have greater influence than the global macroeconomic indicators on the Indian stock index BSE SENSEX in the long run.

Alamsyah and Permana [10] evaluated the effect of five macroeconomic indicators, namely, GDP, investment, import, export, and total expenditure on the Indonesian stock market. They concluded that it is possible to predict the stock market using macroeconomic indicators.

Peng and Li [11] proposed that real GDP, inflation, and unemployment are independent of each other, and these macroeconomic indicators can be used to predict the stock market fairly efficiently.

Shreya and Sangeetha [12] examined the effect of ten macroeconomic indicators namely, gold prices, industrial production, silver prices, exchange rate, interest rate, money supply, oil prices, inflation, forex reserve, and trade balance on the Indian stock market cnx NIFTY. They found a strong correlation between different macroeconomic indicators.

Misra [13] studied the influence of industrial production index, inflation, rate of interest, price of gold, rate of exchange, FII, and the supply of money in the Indian stock market. They identified a long term relationship between macroeconomic indicators and the BSE SENSEX, they also found that inflation and the money supply have no short-term influence on the BSE SENSEX.

Ntshangase *et al.* [14] found that there exists a strong influence between the macroeconomic factors namely,inflation, government expenditure, money supply, interest rate, and exchange rate and the South African stock market.

Omodero and Mlanga [15] found that the interest and exchange rates do not influence the Nigerian stock market. They also found that inflation negatively influences the stock market and GDP positively influences the Nigerian stock market.

Wang *et al.* [16] studied the effect of four macroeconomic indicators, namely, industrial production index, consumer price index, unemployment rate, and federal funds rate on the stock market in the United States of America. They found that these macroeconomic indicators have a very strong influence on the stock market of USA.

Li [17] proved that macroeconomic indicators GDP, CPI, and composite stock index influence the housing prices in China; thus, housing prices and the stock index are interrelated.

Megaravall and Sampagnaro [18] considered the exchange rate and CPI of India, China, and Japan and found their influence on developing economies India and China and developed economy Japan. The results indicate that the exchange rate and CPI influence these three economies equally. The CPI is a substitute for inflation that influences all global economies because it is a producer consumer problem.

Rahman [19] established that there is no relationship between the export of Bangladesh with the macroeconomic factors interest rate, exchange rate, industrial production index, and broad money. She concluded that there are other

extraneous macroeconomic indicators that affect the export of Bangladesh.

Giri and Joshi [20] examined the influence of GDP, inflation, exchange rate, and crude oil price on the Indian stock market. Their result indicates that GDP, inflation, and exchange rate have a positive influence on the stock market, while the crude oil price has a negative influence on the Indian stock market.

Kvietkauskienė and Plakys [21] surveyed the effects of macroeconomic factors on stock market indicators. Through survey they found that GDP, 10-year government benchmark yields, pension funds market size, investment funds market size, employment rate, economic sentiment indicator, purchasing managers indicator, and ZEW have positive influence on the stock market, while gold price, silver price, oil price, country credit rating, unemployment rate, and P/E ratio have negative influence on the stock market. In addition, the consumer price index, inflation, government debt level, and exchange rate have a mixed influence.

Ho and Odhiambo [22] empirically studied the impact of trade openness, the banking sector, and exchange rate on the Philippines stock market. They found that trade openness has a negative impact, while banking sector development and exchange rate have a positive impact on the Philippines stock market.

Thorbecke [23] found that oil price has a strong relationship with the USA stock market. He concluded that after 2010 due to soaring oil production by the USA, the oil prices are beneficial to the USA stock market.

Garg and Kalra [24] studied the effect of six macroeconomic indicators, namely, GDP, gold prices, exchange rate, foreign exchange rate, average inflation, and unemployment rate on the Indian stock market. They found that GDP, gold prices, exchange rate, and foreign exchange rate have a positive impact, and average inflation and unemployment rate have a negative influence on the Indian stock market.

Ho [25] studied of the influence of macroeconomic indicators on the South African stock market. He found that trade openness and interest rate have a negative impact, while banking sector development, economic growth, and exchange rate have a positive impact on the South African stock market.

Lee and Ryu [26] found that the macroeconomic indicators do not affect when analyzed with the linear model, while all the macroeconomic indicators influence the stock market when an analysis was performed using a nonlinear model.

Nijam *et al.* [27] empirically proved that GDP, exchange rate, and interest rate has a positive impact. Inflation has a negative impact, and balance of payment has no impact on Sri Lankan stock exchange.

Heidari *et al.* [28] considered the effect of inflation in health care, oil price shocks, and monetary policy on Iranian pharmaceutical companies. They found that these macroeconomic factors influence the Iranian stock market.

Khan and Khan [29] analyzed the Karachi stock exchange of Pakistan and found that money supply, exchange rate, and interest rate have a strong influence on the stock market in the long run. They also found that exchange rate has a negative impact in the short term.

Ahmed *et al.* [30] studied the effect of three macroeconomic indicators, namely, inflation, interest rate, and exchange rate on Karachi stock exchange-100 and concluded that these macroeconomic indicators were insignificant and does not influence the stock market in Pakistan.

Abbas *et al.* [31] analyzed the influence of macroeconomic factors on stock exchanges in the USA, UK, France, Italy, Canada, Japan, and Germany. They identified that industrial growth and oil prices as the most significant influencing factors affecting the stock markets of the G7 nations.

Ndlovu *et al.* [32] explored the effect of macroeconomic indicators on the Johannesburg stock exchange. They concluded that inflation, interest rate, and money supply have a positive impact on stock prices, while, the exchange rate influences stock prices negatively.

Lee and Brahmasrene [33] studied the effect of macroeconomic factors on the Korean stock market. They concluded that the stock markets are affected by macroeconomic indicators in the long and short term. The influencers were industrial production index, exchange rate, and inflation. They found that money supply does not affect the Korean stock market.

Chowdhury *et al.* [34] found that changes in money supply and government expenditure affect liquidity in emerging Asian stock markets.

Huong [35] identified that inflation has positive influence and corporate tax rate has a negative influence on the corporate capital structure of the stock market in Vietnam.

The result of these studies vary from macroeconomic indicators not influencing the stock market performance to the very strong influence on performance. From this literature review, it is evident that the selection of macroeconomic indicators greatly influences whether any given indicator will, in turn, affect the stock market. The reviewed research has considered a one to ten macroeconomic indicators and studies their effect on stock market indicators. The influence of macroeconomic indicators on stock market in the USA, UK, France, Germany, Italy, China, Japan, South Africa, Nigeria, Vietnam, Canada, Pakistan, Bangladesh, Malyasia, Philippines, Singapore, and India were explored. The selection of the macroeconomic indicator varies from nation to nation, and depends on how well the nation is connected economically with the rest of the world. The following methods were employed in most of the reviewed papers

- Auto Regressive Distributed Lag (ARDL)
- ADF test
- Multi-regression
- Granger causality test

They identify whether there is exogenous variable that affects the stock market through these tests. In some of the cases, the result suggest exogenous variables do affect the stock market performance.

III. METHODOLOGY

The premium dataset trading economics with code SGE offered through quandl.com is considered for analysis, as well as data from the World Gold Council(WGC). Note that, the missing values were substituted by their means in a data cleaning process. The data was present in different granularity. For example, GDP was recorded yearly, change in inventories was recorded quarterly, the balance of trade was recorded monthly, foreign exchange reserve was recorded weekly, and treasury bill 91 days yield was recorded daily. All the macroeconomic indicators are transformed into monthly observations. The forty-four macroeconomic indicators for a period of eight years from January 2011 to December 2018 were considered. Thus, the dataset consists of forty-four columns and ninety-six rows.

A Correlation matrix of all the forty-four macroeconomic indicators was found. KMO and Bartlett tests were applied to identify the suitability of applying a dimensionality reduction algorithm [36]. The Bartlett B and KMO K values were calculated using equations 2 and 3 respectively. The PCA algorithm was applied [37], and the number of factors to be retained was obtained by applying the K1-Kaiser method and scree test [38]. According to the K1-Kaiser method all factors whose eigenvalue is greater than 1 are retained. In a scree plot, if there is a break between two factors, then those factors are retained. In addition, in a scree plot, user select the threshold. After identifying the number of factors, PCA with varimax rotation was utilized to identify the belongingness of each macroeconomic indicator to each factors. In addition, ADF test was applied to identify the autocorrelation of prominent macroeconomic indicators and stock indices. It was observed that both endogenous and exogenous variables were initially non-stationary and the residual was stationary. The endogenous and exogenous variables are stationary at first difference which implies that there is cointegration between them; thus VECM should be performed. VECM was performed, and correction was applied. Regression analysis was applied with NSE Nifty as the dependent variable and prominent macroeconomic indicators from each of the factors as independent variables. Regression analysis was repeated with BSE SENSEX as the dependent variable. The purpose of applying regression analysis was to identify which factors influence the stock market positively and which factors have a negative impact. The Figure.2. depicts the block diagram of the process. The approach is similar to that of KDD (Figure1).

Finally, a feedforward neural network with backpropagation with sigmoid and ReLU as activation functions were applied to predict the stock index prices from prominent macroeconomic indicators belonging to each of the factors.

IV. RESULTS

The forty-four macroeconomic indicators considered are listed in Table1. The partial correlation matrix obtained for the macroeconomic indicators is listed in Table2. To check whether dimensionality reduction can be applied is done through KMO and Bartlett test. If some variables can collectively imply other variables, then the KMO test cannot be applied. In such cases, one must drop dependent

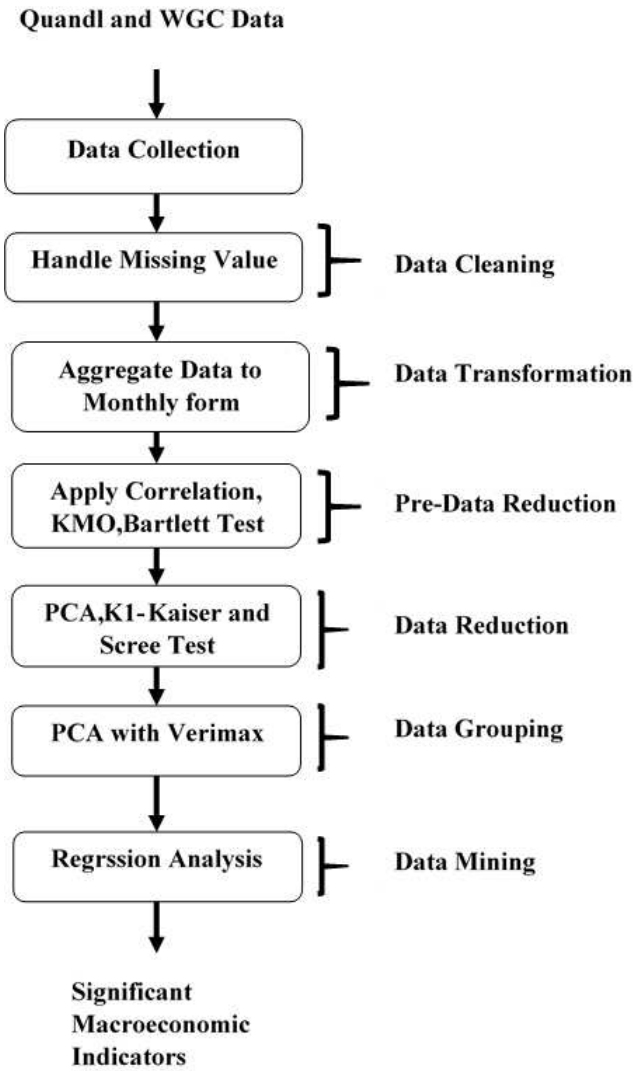

FIGURE 2. Block diagram of methodology.

TABLE 1. List of macroeconomic indicators.

The Macroeconomic Indicators	
BALANCE-OF-TRADE	CAR SALES
COMPETITIVE INDEX	CURRENT ACCOUNT
CRUDE OIL PRODUCTION	EXPORT
INDUSTRIAL PRODUCTION	FOOD INFLATION
INFLATION RATE	FDI
INFRASTRUCTURE OUTPUT	IMPORT
MANUFACTURING PMI	REMITTANCES
PRIME LENDING RATE	VAT
TOTAL EXTERNAL DEBT	WHOLE SALE INDEX SALE
EXPORT PRICES	GOLD PRICES
MONEY SUPPLY M1	MONEY SUPPLY M2
MONEY SUPPLY M3	PRODUCER PRICES
CENTRAL-BANK-BALANCE-SHEET	CENTRAL-GOVT-BUDGET
CHANGE-IN-INVENTORIES	FOREX
CONSUMER PRICE INDEX	GDP
TREASURY BILL 91 DAYS YIELD	GOVT SPENDING
CURRENT-ACCOUNT-TO-GDP	GOVT BUDGET VALUE
GOVT BOND 10 YEARS YIELD	SERVICE PMI
GROSS FIXED CAPITAL FORMATION	TOURIST ARRIVALS
MANUFACTURING PRODUCTION	CORPORATE TAX RATE
TOTAL DISPOSABLE INCOME	UNEMPLOYMENT RATE
GOLD RESERVE	INTEREST RATE

TABLE 2. The correlation of balance-of-trade with other macroeconomic indicators.

	BALANCE-OF-TRADE
BALANCE-OF-TRADE	1
CAR SALES	0.032689557
CENTRAL-BANK-BALANCE-SHEET	0.186131557
CENTRAL-GOVT-BUDGET	0.338523302
CHANGE-IN-INVENTORIES	0.454132671
Competitive Index	-0.276301522
Consumer Price Index	0.183039341
Corporate Tax Rate	0.165509306
CRUDE OIL PRODUCTION	0.028258073
Current Account	0.824001414
CurrentAccount-To-GDP	0.52776772
Export	-0.219090587
Export Prices	0.206546739
FOOD INFLATION	-0.07100893
FDI	0.138589187
Foreign Exchange Reserve	-0.08392864
GDP	0.036090317
Gold Prices	-0.473007161
Gold Reserve	-0.229922699
GOVT SPENDING	-0.079940689
Govt Bond 10 Years	-0.127055361
Govt Budget Value	-0.088137462
Gross Fixed Capital Formation	-0.09887271
Import	-0.865190182
Industrial Production	0.18052523
Inflation Rate	-0.175680471
Infrastructure Output	-0.009809346
Interest Rate	-0.214179898
Manufacturing PMI	-0.322361362
Manufacturing Production	0.074896552
Money Supply M1	0.090149284
Money Supply M2	0.077493738
Money Supply M3	0.131569445
Prime Lending Rate	-0.314670875
Producer Prices	0.006939405
REMITTANCES	-0.087347149
VAT	0.039028954
SERVICE PMI	-0.336477072
Total Disposable Personal Income	0.106999895
Total External Debt	0.178491233
Tourist Arrivals	0.081526841
Treasury Bill 91 Days Yield	-0.235444216
Unemployment Rate	-0.313816984
Whole Sale Price Index Sale	-0.421584221

TABLE 3. Result of the bartlett test.

Bartlett Test	p-value
13133.114641070972	0.0

variables from the dataset. So to conduct the KMO test, the two macroeconomic indicators, namely, central-government-budget which depends on government budget value and competitive index which depends on balance-of-trade, are dropped. The Bartlett test is conducted with all the forty-four macroeconomic indicators as it does not have the constraint of dependency between variables. The result of the Bartlett test is presented in Table3. As p-value is zero for the Bartlett test, which is less than 0.05, and KMO value is 0.79, which is greater than 0.5, one can consider reducing the dimensionality of the problem.

The PCA algorithm results are shown in Table4. According to the K1-Kaiser method, seven factors whose eigenvalue was greater than one were considered. The number of factors identified through the scree plot matches that obtained by the K1-Kaiser method. The scree plot is depicted in Figure.10. PCA with varimax rotation was applied with seven factors. The PCA with varimax gives the belongingness of each of the macroeconomic indicator for each factor. Table5 depicts the belongingness of prominent macroeconomic indicators in each of the seven factors obtained from PCA with varimax. The macroeconomic indicators fall to the factor for which its belongingness is highest. The macroeconomic indicators under each of the factors are given in Tables7,8,9,10,11,12,and 13. The seven factors can explain 80% of the variance among the macroeconomic indicators. Table6 depicts the cumulative variance explained by each factor. The prominent macroeconomic indicators from each of the factors were considered for regression analysis. A list of prominent and representative macroeconomic indicators is given in Table14.

TABLE 4. The eigenvalues.

	Original_Eigenvalues
0	20.67582944
1	5.866897157
2	2.773958363
3	2.61905755
4	1.910523321
5	1.483256573
6	1.191086054
7	0.956694212
8	0.854180419
9	0.734561854
10	0.700021178
11	0.604315062
12	0.581189611
13	0.461456232
14	0.386889878
15	0.348617089
16	0.285537268
17	0.258298647
18	0.182184826
19	0.166529597
20	0.13685149
21	0.132047959
22	0.112468817
23	0.09117694
24	0.088674186
25	0.071917919
26	0.063940802
27	0.060119785
28	0.053982342
29	0.048660562
30	0.027592406
31	0.020528665
32	0.014110186
33	0.012535432
34	0.010379751
35	0.005428995
36	0.004463487
37	0.001893559
38	0.001390356
39	0.000477898
40	0.000272324
41	1.80557E-06
42	1.94535E-15
43	-1.5735E-15

TABLE 5. Belongingness of prominent macroeconomic indicators to each of the factors.

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
GDP	0.972948	0.020884	0.051597	-0.09892	0.106961	0.087275	-0.081091
IMPORT	0.019677	0.934982	0.024314	-0.03348	-0.18479	-0.00678	0.227413
INDUSTRIAL PRODUCTION	0.077752	-0.01995	0.990015	0.071127	0.076658	0.032767	-0.02155
CORPORATE TAX RATE	0.017527	-0.12688	0.078537	0.91941	0.08866	-0.02634	0.024358
CURRENT ACCOUNT-TO-GDP	0.102613	-0.51561	0.060485	0.283664	0.873045	0.046151	0.257651
TOURIST ARRIVALS	0.605483	-0.04064	0.038181	-0.02347	0.043013	0.861472	-0.09876
PRIME LENDING RATE	-0.59731	0.139213	-0.2293	-0.02538	-0.3383	-0.03808	0.933821

TABLE 6. Cumulative variance of each factor.

	Factor1	Factor2	...	Factor6	Factor7
SS Loadings	18.937795	5.096869	...	1.646201	1.233219
Proportion Var	0.430404	0.115838	...	0.037414	0.033028
Cumulative Var	0.430404	0.546242	...	0.767201	0.800229

TABLE 7. The macroeconomic indicators belonging to factor-1.

Factor 1
CAR SALES
CENTRAL-BANK-BALANCE-SHEET
CENTRAL-GOVT-BUDGET
COMPETITIVE INDEX
CONSUMER PRICE INDEX
EXPORT PRICES
FDI
FOREX
GDP
GOLD RESERVE
GOVT SPENDING
GOVT BUDGET VALUE
GROSS FIXED CAPITAL FORMATION
MONEY SUPPLY M1
MONEY SUPPLY M2
MONEY SUPPLY M3
PRODUCER PRICES
REMITTANCES
VAT
TOTAL DISPOSABLE INCOME
TOTAL EXTERNAL DEBT
UNEMPLOYMENT RATE

TABLE 8. The macroeconomic indicators belonging to factor-2.

Factor 2
EXPORT
FOOD INFLATION
GOLD PRICES
IMPORT
INFLATION RATE
MANUFACTURING PMI
WHOLE SALE PRICE INDEX

Regression analysis was performed once for the NSE Nifty index and once for the BSE SENSEX index. The regression statistics results are depicted in Tables15 and 16 respectively. The ANOVA result are presented in Tables17 and 18 respectively. The regression equation for the NSE Nifty index is depicted in Equation-4, and the regression equation for BSE SENSEX is presented in Equation-5.

TABLE 9. The macroeconomic indicators belonging to factor-3.

Factor 3
INDUSTRIAL PRODUCTION
INFRASTRUCTURE OUTPUT
MANUFACTURING PRODUCTION
SERVICE PMI

TABLE 10. The macroeconomic indicators belonging to factor-4.

Factor 4
CORPORATE TAX RATE
GOVT BOND 10 YEARS

TABLE 11. The macroeconomic indicators belonging to factor-5.

Factor 5
BALANCE-OF-TRADE
CURRENT ACCOUNT
CURRENT ACCOUNT-TO-GDP

TABLE 12. The macroeconomic indicators belonging to factor-6.

Factor 6
TOURIST ARRIVALS

TABLE 13. The macroeconomic indicators belonging to factor-7.

Factor 7
CHANGE-IN-INVENTORIES
CRUDE OIL PRODUCTION
INTEREST RATE
PRIME LENDING RATE
TREASURY BILL 91 DAYS YIELD

$$
\begin{aligned}
NIFTY &= -7951.43579275191 \\
&\quad + 159.713597784824 * CORPORATETAXRATE \\
&\quad + 214.932823343983 * CURRENTACCOUNTTOGDP \\
&\quad + 4.93315919740711 * GDP \\
&\quad + 0.0346532627892473 * IMPORT
\end{aligned}
$$

TABLE 14. The prominent macroeconomic indicator belonging to each category.

FACTOR NUMBER	FACTOR NAME
FACTOR 1	GDP
FACTOR 2	IMPORT
FACTOR 3	INDUSTRIAL PRODUCTION
FACTOR 4	CORPORATE TAX RATE
FACTOR 5	CURRENT ACCOUNT TO GDP
FACTOR 6	TOURIST ARRIVALS
FACTOR 7	PRIME LENDING RATE

TABLE 15. The regression statistics for NSE NIFTY index.

Regression Statistics	
Multiple R	0.961408563
R Square	0.924306425
Adjusted R Square	0.918285345
Standard Error	549.7053518
Observations	96

TABLE 16. The regression statistics for BSE SENSEX index.

Regression Statistics	
Multiple R	0.961883412
R Square	0.925219697
Adjusted R Square	0.919271264
Standard Error	1754.668984
Observations	96

$$
\begin{aligned}
&- 14.3079912478492 * INDUSTRIALPRODUCTION \\
&- 118.4865379 * PRIMELENDINGRATE \\
&- 6.56555E - 05 * TOURISTARRIVALS \qquad (4)
\end{aligned}
$$

$$
\begin{aligned}
SENSEX &= -52808.87195 \\
&+ 748.3161561 * CORPORATETAXRATE \\
&+ 380.2482201 * CURRENTACCOUNTTOGDP \\
&+ 18.88783164 * GDP \\
&- 0.063282173 * IMPORT \\
&- 31.0047566 * INDUSTRIALPRODUCTION \\
&+ 1683.365331 * PRIMELENDINGRATE \\
&- 0.002197158 * TOURISTARRIVALS \qquad (5)
\end{aligned}
$$

It is evident from the results that even within the nation, different indices are affected by different macroeconomic indicators differently. From equation4, it is evident that the macroeconomic indicators belonging to factors 1, 2, 4, and 5 have a positive relationship with the NSE Nifty index, and the macroeconomic indicators belonging to factors 3, 6, and 7 have a negative relation with NSE Nifty. From Equation 5 it is clear that macroeconomic indicators belonging to factors 1, 4, 5, and 7 have a positive relationship with the BSE SENSEX while macroeconomic indicators belonging to factors 2, 3, and 6 have a negative relation.

TABLE 17. The ANOVA test results for NSE Nifty index.

ANOVA test results					
	df	SS	MS	F	Significance F
Regression	7	324712911.8	46387558.83	153.5117377	1.67649E-46
Residual	88	26591485.7	302175.9738		
Total	95	351304397.5			

TABLE 18. The ANOVA test results for BSE SENSEX index.

ANOVA test results					
	df	SS	MS	F	Significance F
Regression	7	3352206189	478886598.4	155.540068	9.85119E-47
Residual	88	270939965.5	3078863.244		
Total	95	3623146154			

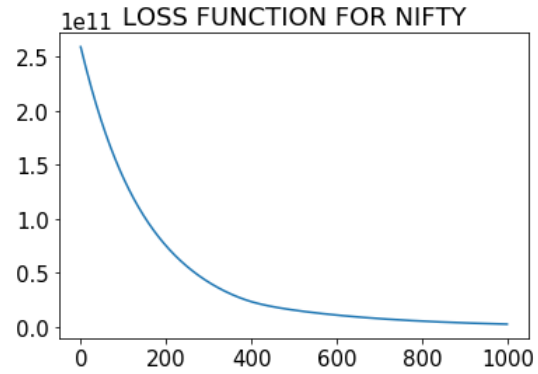

FIGURE 3. Loss function of ANN with ReLU for Nifty as the target variable.

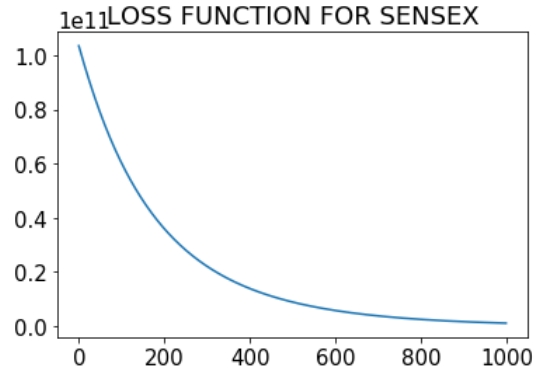

FIGURE 4. Loss function of ANN with ReLU for SENSEX as the target variable.

A nonlinear ANN with the following configuration was constructed:

1) ReLu activation function
2) Alpha value of 0.0001
3) Beta-1 value of 0.9; Beta-2 value of 0.999
4) Early stopping was set to false
5) Epsilon value of 1e-08
6) Maxiteration set to 1000
7) Validation fraction was set to 10%

The loss function of the feedforward neural network with backpropagation and ReLu activation function for the Nifty index as the output variable is shown in Figure.3, and the same

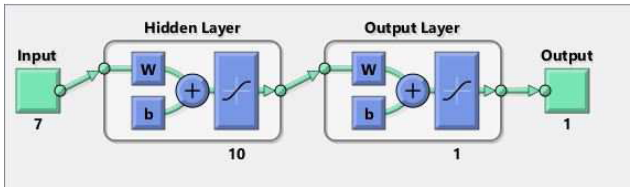

FIGURE 5. ANN with three layers.

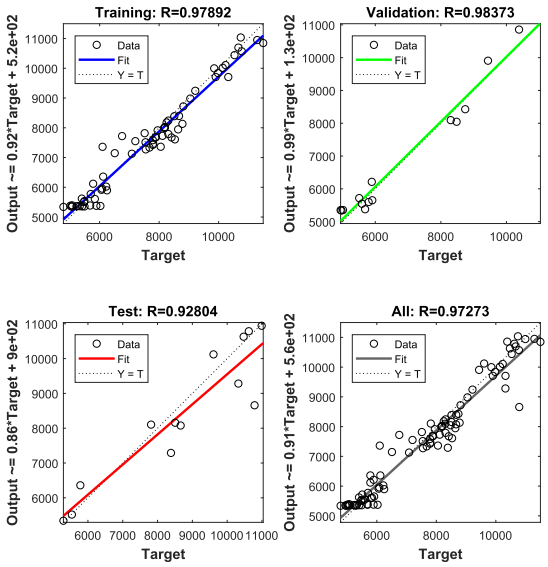

FIGURE 6. R plot for ANN with Nifty as target variable.

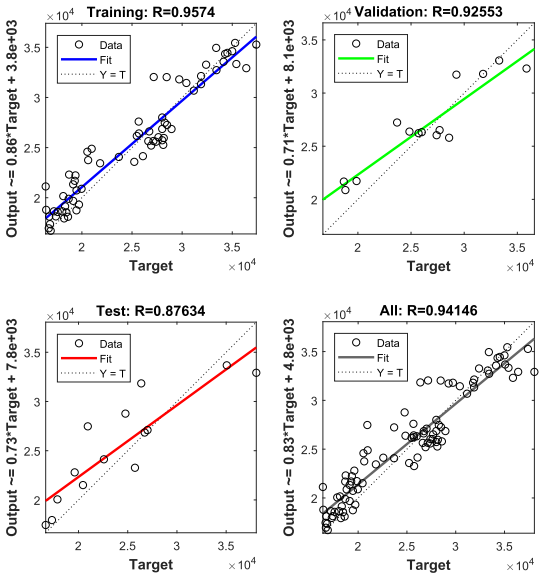

FIGURE 7. R plot for ANN with SENSEX as target variable.

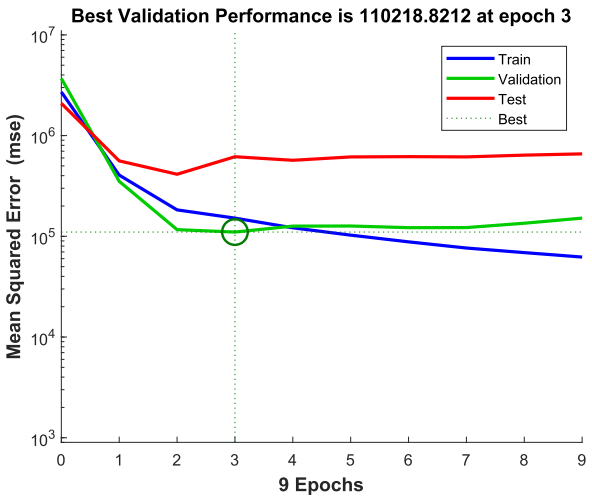

FIGURE 8. Performance plot for ANN with nifty as target variable.

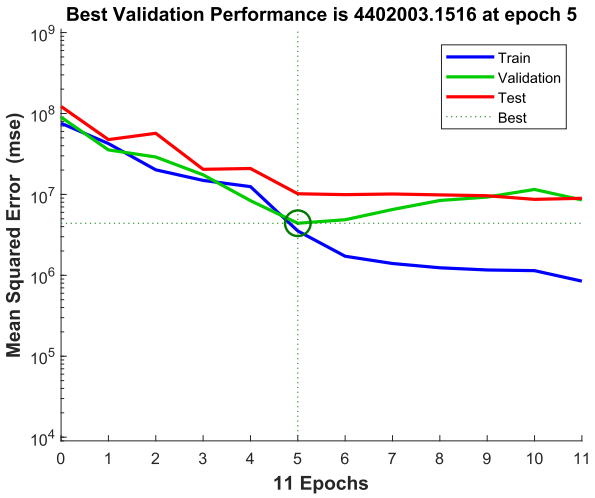

FIGURE 9. Performance plot for ANN with SENSEX as target variable.

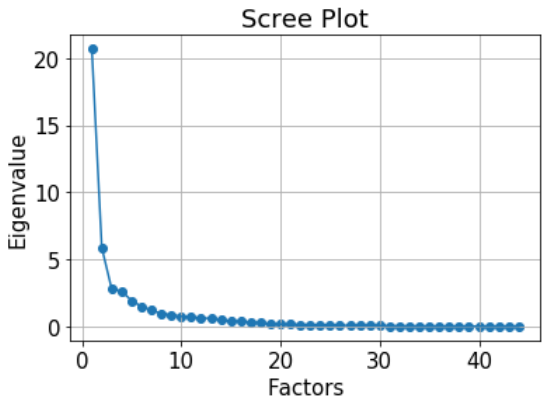

FIGURE 10. The scree plot.

for the SENSEX as the output variable is shown in Figure.4. It is observed that with the ReLU activation function the optimizer would not converge even after thousand epochs.

The feedforward neural network with backpropagation and sigmoid activation function demonstrated better performance and accuracy. The feedforward neural network was constructed with three layers. There were seven neurons in the input layer corresponding to the seven factors. The optimal number of hidden layer neurons was obtained by trial and error. Ten neurons were present in the hidden layer, and the

output layer consists of one neuron corresponding to the stock market index. During the training phase, 60% of the data was allotted for training, 20% for validation, and 20% for testing. The built ANN is depicted in Figure.5. R plots for ANN with Nifty and SENSEX as target variables are depicted in Figures.6 and 7, respectively. The performance plots for Nifty and SENSEX as the target variables are presented in Figures.8 and 9, respectively.

The test data span from January 2019 to August 2019. After testing the ANN for test data, 92% accuracy was obtained with NSE Nifty as the output variable, and 87% accuracy was obtained with BSE SENSEX as the output variable.

V. CONCLUSION AND FUTURE WORK

This paper provides a scientific way of identifying significant macroeconomic indicators influencing stock markets. The study can be applied to any global stock market. The case of Indian stock markets NSE Nifty and BSE SENSEX was taken as sample indices to demonstrate the methodology. KMO and Bartlett tests substantiated the necessity of dimensionality reduction. PCA algorithm was applied on the macroeconomic indicators, and with K1-Kaiser and scree plot, seven factors were identified. PCA with varimax rotation explained approximately 80% of the variance among the macroeconomic indicators. The ADF test proved the existence of cointegration between the exogenous and endogenous variables. The regression equation uncovered the fact that within a nation, different stock exchanges are influenced by different macroeconomic indicators differently. ANNs with ReLu and Sigmoid activation functions were employed to predict the stock market movement using the prominent macroeconomic indicators from each factor. An accuracy of 92% in case of Nifty and 87% in case of SENSEX were obtained. Even ANN results prove that macroeconomic indicators affect different stock markets differently with the same nation. The proposed methodology can be applied to the stock market of any nation to identify the significant financial indicators influencing it. In future, this work can be extended to identify qualitative data influencing stock markets.

ACKNOWLEDGMENT

The authors would like to thank Enago (www.enago.com) for the English language review.

REFERENCES

[1] J. Shlens, "A tutorial on principal component analysis," 2014, *arXiv:1404.1100*. [Online]. Available: https://arxiv.org/abs/1404.1100

[2] T. J. Flavin and M. R. Wickens, "Macroeconomic influences on optimal asset allocation," *Rev. Financial Econ.*, vol. 12, no. 2, pp. 207–231, 2003.

[3] M. N. Zarina, "Asset pricing and macro factors in ASEAN5," in *Proc. IEEE Colloq. Humanities, Sci. Eng.*, Dec. 2011, pp. 592–596.

[4] X. Chen and X. Jin "Detecting the macroeconomic factors in Chinese stock market returns: A generalized dynamic factor model approach," in *Proc. 2nd IEEE Int. Conf. Inf. Financial Eng.*, Sep. 2010, pp. 184–188.

[5] S. Sudhakaran and P. A. S. Balasubramanian, "A study on the impact of macroeconomic factors on S&P BSE Bankex returns," in *Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, Sep. 2016, pp. 2614–2618.

[6] S. Claessens and M. A. Kose, *Asset Prices and Macroeconomic Outcomes: A Survey*. Washington, DC, USA: World Bank, 2017.

[7] Q. J. Zhang, P. Hopkins, S. Satchell, and R. Schwob, "The link between macro-economic factors and style returns," *J. Asset Manage.*, vol. 10, no. 5, pp. 338–355, 2009.

[8] K. Gurloveleen and B. S. Bhatia, "An impact of macroeconomic variables on the functioning of Indian stock market: A study of manufacturing firms of BSE 500," *J. Stock Forex Trading*, vol. 5, no. 1, pp. 160–167, 2015.

[9] C. Parmar, "Empirical relationship among various macroeconomics variables on Indian stock market," *Int. J. Adv. Res. Comput. Sci. Manage. Stud.*, vol. 1, no. 6, pp. 190–197, 2013.

[10] A. Alamsyah and M. F. Permana, "Artificial neural network for predicting indonesian economic growth using macroeconomics indicators," in *Proc. Int. Symp. Adv. Intell. Inform. (SAIN)*, Aug. 2018, pp. 15–19.

[11] Z. Peng and X. Li "Application of a multi-factor linear regression model for stock portfolio optimization," in *Proc. Int. Conf. Virtual Reality Intell. Syst. (ICVRIS)*, Aug. 2018, pp. 367–370.

[12] S. Agrawal and R. Sangeetha, "The impact of macroeconomic determinants on the performance of the Indian stock market," *Int. J. Bus. Manage. Invention*, vol. 8, no. 3, pp. 16–23, 2019.

[13] P. Misra, "An investigation of the macroeconomic factors affecting the Indian stock market," *Australas. Accounting, Bus. Finance J.*, vol. 12, no. 2, pp. 71–86, 2018.

[14] K. Ntshangase, K. F. Mingiri, and M. M. Palesa, "The interaction between the stock market and macroeconomic policy variables in South Africa," *J. Econ.*, vol. 7, no. 1, pp. 1–20, 2016.

[15] C. O. Omodero and S. Mlanga, "Evaluation of the impact of macroeconomic variables on stock market performance in Nigeria," *Bus. Manage. Stud.*, vol. 5, no. 2, pp. 34–44, 2019.

[16] Y. Wang, H. Lu, and S. Xiong, "An empirical analysis of macroeconomic factors and stock returns," in *Proc. Int. Conf. Data Mining, Commun. Inf. Technol. (DMCIT)*, New York, NY, USA, 2017, Art. no. 25.

[17] P. Li, "Macroeconomic determinants of China's housing prices," in *Proc. 8th Int. Conf. E-Bus., Manage. Econ. (ICEME)*, New York, NY, USA, 2017, pp. 102–106.

[18] A. V. Megaravalli, G. Sampagnaro, and L. Murray, "Macroeconomic indicators and their impact on stock markets in ASIAN 3: A pooled mean group approach," *Cogent Econ. Finance*, vol. 6, no. 1, 2018, Art. no. 1432450.

[19] F. Rahman, "Impact of selected macro-economic variables on the export performance of Bangladesh," *IIOSR J. Econ. Finance*, vol. 8, no. 2, pp. 21–27, 2017.

[20] A. K. Giri and P. Joshi, "The impact of macroeconomic indicators on Indian stock prices: An empirical analysis," *Stud. Bus. Econ.*, vol. 12, no. 1, pp. 61–78, 2017.

[21] E. A. Kvietkauskien and M. Plakys, "Impact indicators for stock markets return," *Poslovna izvrsnost: Znanstveni časopis za promicanje kulture kvalitete i poslovne izvrsnosti*, vol. 11, no. 2, pp. 59–83, 2017.

[22] S.-Y. Ho, N. M. Odhiambo, and D. McMillan, "Analysing the macroeconomic drivers of stock market development in the Philippines," *Cogent Econ. Finance*, vol. 6, no. 1, 2018, Art. no. 1451265.

[23] W. Thorbecke, "Oil prices and the U.S. Economy: Evidence from the stock market," *J. Macroecon.*, vol. 61, Sep. 2019, Art. no. 103137.

[24] K. Garg and R. Kalra, "Impact of macroeconomic factors on Indian stock market," *KIIT-J. Manage.*, vol. 14, no. 1, pp. 134–145, 2018.

[25] S.-Y. Ho, "Macroeconomic determinants of stock market development in South Africa," *Int. J. Emerg. Markets*, vol. 14, no. 2, pp. 322–342, 2019.

[26] G. Lee and D. Ryu, "Asymmetry in the stock price response to macroeconomic shocks: Evidence from the Korean market," *J. Bus. Econ. Manage.*, vol. 19, no. 2, pp. 343–359, 2018.

[27] H. M. Nijam, S. M. Ismail, and A. M. Musthafa, "The impact of macro-economic variables on stock market performance; evidence from Sri Lanka," *J. Emerg. Trends Econ. Manage. Sci.*, vol. 9, no. 2, pp. 115–121, 2018.

[28] H. Heidari, A. R. Kahriz, and Y. Mohammadzadeh, "Stock market behavior of pharmaceutical industry in Iran and macroeconomic factors," *Econ. Change Restructuring*, vol. 52, no. 3, pp. 255–277, 2019.

[29] J. Khan and I. Khan, "The impact of macroeconomic variables on stock prices: A case study of Karachi stock exchange," *J. Econ. Sustain. Develop.*, vol. 9, no. 13, pp. 15–25, 2018.

[30] W. Ahmed, A. Aslam, and A. Hakeem, "The influence of macro-economic factors on stock prices of KSE 100 index," *Int. J. Res. Sci. Invention*, vol. 5, no. 11, pp. 35–44, 2018.

[31] G. Abbas, D. G. McMillan, and S. Wang, "Conditional volatility nexus between stock markets and macroeconomic variables: Empirical evidence of G-7 countries," *J. Econ. Stud.*, vol. 45, no. 1, pp. 77–99, 2018.
[32] B. Ndlovu, F. Faisa, N. G. Resatoglu, and T. Türsoy, "The impact macroeconomic variables on stock returns: A case of the johannesburg stock exchange," *Romanian Stat. Rev.*, vol. 2, pp. 1–18, Jun. 2018.
[33] J. W. Lee and T. Brahmasrene, "An exploration of dynamical relationships between macroeconomic variables and stock prices in korea," *J. Asian Finance, Econ. Bus.*, vol. 5, no. 3, pp. 7–17, 2018.
[34] A. Chowdhury, M. Uddin, and K. Anderson, "Liquidity and macroeconomic management in emerging markets," *Emerg. Markets Rev.*, vol. 34, pp. 1–24, Mar. 2018.
[35] P. T. Q. Huong, "Macroeconomic factors and corporate capital structure: Evidence from listed joint stock companies in Vietnam," *Int. J. Financial Res.*, vol. 9, no. 1, pp. 31–40, 2018.
[36] T. Venkatesan and R. Venkataraman, "Analysis of factors determining financial literacy using structural equation modelling," *SDMIMD J. Manage.*, vol. 9, no. 2, pp. 19–29, 2018.
[37] F. Hayat, M. D. S. Pirzada, and A. A. Khan, "The validation of granger causality through formulation and use of finance-growth-energy indexes," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1859–1867, Jan. 2018.
[38] T. B. Amene, "Obstacles of women entrepreneurial success in the case of Chiro City," *Int. J. Develop. Res.*, vol. 8, no. 5, pp. 20613–20621, 2018.

PRAKASH K. AITHAL is currently pursuing the Ph.D. degree with the Manipal Academy of Higher Education, Manipal, India, where he is also an Assistant Professor with the Department of Computer Science and Engineering, Manipal Institute of Technology. He has presented several articles in national and international conferences. His work has been published in several international journals. His current research interest includes data analytics in financial domain.

DINESH ACHARYA U. received the Ph.D. degree from the Manipal Academy of Higher Education, Manipal, India, in 2008, where he is currently a Professor with the Department of Computer Science and Engineering, Manipal Institute of Technology. He has presented several articles in national and international conferences. His work has been published in several international journals. His current research interests include data analytics in healthcare, agriculture, and financial sectors.

GEETHA M. received the Ph.D. degree from NITK, Surathkal, in 2010. She is currently a Professor with the Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India. She has presented several articles in national and international conferences. Her work has been published in several international journals. Her current research interests include data mining, text mining in healthcare, and financial sectors.