A STUDY ON INDIAN STOCK MARKET MODELING USING ARTIFICIAL NEURAL NETWORKS

Thesis

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

By

ARJUN R

SCHOOL OF MANAGEMENT

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA,

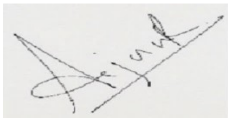
SURATHKAL, MANGALORE – 575 025

AUGUST, 2021

D E C L A R A T I O N

By the Ph.D. Research Scholar

I hereby *declare* that the Research Thesis entitled **A Study on Indian Stock Market Modeling using Artificial Neural Networks,** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfilment of the requirements for the award of the Degree of **Doctor of Philosophy in Management** (specialized in Information Systems) is a *bonafide report of the research work carried out by me*. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.



Arjun R

Reg. No. 1551525HM15F01

School of Management

Place: NITK, Surathkal

Date: 10-06-2021

CERTIFICATE

This is to *certify* that the Research Thesis entitled **A Study on Indian Stock Market Modeling using Artificial Neural Networks** submitted by Arjun R, (Register Number: 1551525HM15F01), as the record of the research work carried out by him, is *accepted as the Research Thesis submission* in partial fulfilment of the requirements for the award of the degree of **Doctor of Philosophy**.

....

Dr. Suprabha K.R.

Research Supervisor

Assistant Professor

School of Management

....

3/8/21

Dr. Pavan Kumar S

Chairman – DRPC

School of Management

NIT Karnataka- 575 025

Dr. S. PAVAN KUMAR

Head, School of Management

National Institute of Technology Karnataka

Post Srinivasnagar, Surathkal, D.K.-575025

ACKNOWLEDGMENT

This incredible journey of research was fulfilled because of the effort and kind-heartedness of many special persons. First and foremost, I express my prayers to the Almighty. Next, I express deep appreciation to my guide Dr. Suprabha K.R for believing in my skills, ability. Also, profoundly thank the department Head, Dr. Pavan Kumar S, for instilling confidence and helping out whenever necessary. The special lectures by experts became a turning point for my work.

Special mention to all faculty members within and outside institute for helping to broaden my knowledge. I also pay my condolences to the late Dr. Sunil D'Souza, who paved my path to research, and also my late friend Mrs. Nandhini Menon. Great cheers, and thanks to all my colleagues at the institute. The kind of friendship and memories shared are so precious. I profusely thank The Director, NITK, Prof. Umamaheshwar Rao K, for generous support through fellowship. I place on record gratitude to Dr. Santhosh Kumar G, CUSAT, all teachers, and Dr. Madhu Kumar S.D. I thank my evergreen pals, Mr. Shine S.S, Mr. Sreekanth B.S, Mr. Unnikrishnan K.C, and others who supported me through thick and thin.

Last but not least, dedicate this achievement to my family. My dear father, Shri. Somanathan E and my mother, Smt. Remadevi S.N has motivated with all moral support. Highly indebted to my wife, Smt. Anjana G Anand and our daughter Anikha for endless love and support. Also, thanks to my sister Smt. Geethu R.S and family.

Arjun R

ABSTRACT

The Indian stock exchange markets, specifically in banking, are dynamic due to diverse micro and macro-level factors. Current research aims to build a predictive model for the banking sector stock market. Statistical estimation models are tested to identify the best predictive parameters. For intelligent decision support design, artificial neural network architectures are simulated. Preliminary results suggest that market volatility has a lesser impact than fundamental and technical indicators, contrary to random walk theory. The artificial neural networks have superior accuracy for National Stock Exchange prediction. However, it requires model retraining, real-time market data, whereas time-series models suit Bombay Stock Exchange forecasting.

Additionally, banking stock performance strongly correlates with technological advancements. Hence, bibliometric analysis extracts areas for the implementation of predictive information systems. An integrated framework is envisaged to adopt blockchain and fintech technologies stimulating organizational impact. Lastly, future research directions provide methodological progress along with the challenges outlined.

Keywords: *Predictive information systems*, *Stock market forecasting*, *Neural networks*, *Banking*, *Business intelligence*

TABLE OF CONTENTS

i. **List of Figures** ... iii

ii. **List of Tables** ... iv

iii. **Nomenclature** ... v

Chapter 1. ... 1

INTRODUCTION ... 1

1.1. **BACKGROUND** ... 2

1.2. **BANKING AND FINTECH** ... 4

1.3. **ROLE OF ANALYTICS AND DSS** ... 5

1.4. **MOTIVATION AND SCOPE** ... 6

1.5. **ORGANIZATION OF THESIS** ... 10

Chapter 2. ... 11

STOCK MARKET MODELING: A CONCEPTUAL FRAMEWORK ... 11

2.1. THEORETICAL PREMISE ... 11

2.2. REVIEW OF RELATED WORK ... 14

2.3. LITERATURE MAP ... 19

2.4. RESEARCH GAPS ... 22

2.5. RESEARCH QUESTIONS ... 24

2.6. RESEARCH OBJECTIVES ... 25

2.7. RESEARCH HYPOTHESES ... 25

Chapter 3 ... 27

RESEARCH METHODOLOGY ... 29

3.1. VARIABLES ... 29

3.1.1. **Price-to-earnings ratio (P/E)** ... 29

3.1.2. **Price-to-book ratio (P/B)** ... 30

3.1.3. **Dividend yield** ... 30

3.1.4. **Beta** ... 31

3.1.5. **Stock index history** ... 31

3.1.6. **Stock price search interest** ... 32

3.1.7. **Closing index** ... 32

3.2. **RESEARCH MODELS** ... 33

3.2.1. Multilayer perceptron models ... 33

3.2.2. Backpropagation (BP) algorithm: ... 34

3.3. SAMPLING DESIGN ... 36

3.4. DATA SOURCES AND TOOLS 36
3.4.1. Simulation and data analysis 37
3.4.2. Model implementations 37
3.4.2.1. ARMA model 37
3.4.2.2. Radial basis function (RBF) model 38
3.4.2.3. Vector Auto-regression (VAR) 38
3.4.2.4. NARX model 39
Chapter 4. 43
ANALYSIS AND DISCUSSION 43
4.1. CHARACTERISTICS OF DATA 43
4.2. EXPERIMENTAL SETTINGS 47
4.2.1. Statistical methods 48
4.2.2. Pre-processing 48
4.3. MODEL PERFORMANCE RESULTS 51
4.4. DISCUSSION 66
4.4.1. Major factors affecting banking index 66
4.4.2. Efficacy of statistical models in prediction 66
4.4.3. Comparative analysis of neural network models 67
4.4.4. An optimum neural network model for stock index prediction 68
4.4.5. Implementation of intelligent decision support systems in the sector 68
Chapter 5. 73
CONCLUSIONS 73
5.1. SUMMARY OF FINDINGS 73
5.2. CONTRIBUTIONS TO BODY OF KNOWLEDGE 75
5.3. IMPLICATIONS AND LIMITATIONS 76
5.3.1. Technical implications 76
5.3.2. Managerial implications 77
5.3.3. Limitations of research 77
5.4. FUTURE DIRECTIONS AND SCOPE 78
TABLE 4.3.7: PREDICTIVE ACCURACY SUMMARY 81
TABLE 4.3.9: GLOBAL STOCK INDICES CORRELATIONS 83
TABLE 4.4.1: TOP IMPACT MAKING 5 STUDIES (SOURCE: SCOPUS) 85
TABLE 4.4.2: TOP IMPACT MAKING 5 STUDIES (SOURCE: WEB OF SCIENCE) . 87
Publications in PhD programme 109

i. List of Figures

FIGURE 2.1.1: FUNDAMENTAL ANALYSIS 11
FIGURE 2.1.2: TECHNICAL ANALYSIS 12

FIGURE 2.2.1: BIBLIOMETRIC COUPLING ANALYSIS 18

FIGURE 2.3.1: STOCK MARKET PREDICTION MODELS 20

FIGURE 2.4.1: TIMELINE OF PAST RESEARCH 23
FIGURE 2.4.2: GLOBAL RESEARCH OUTPUT 23

FIGURE 3.2.1: A GENERIC NEURAL NETWORK 34
FIGURE 3.2.2: RESEARCH DESIGN 35

FIGURE 4.2.1: BSE BANKEX CLOSING & CYCLICAL COMPONENT 47
FIGURE 4.2.2: BSE BANKEX CLOSING DATA WITH INDICATORS 47
FIGURE 4.2.3: NSE NIFTY BANK DATA WITH INDICATORS 48
FIGURE 4.2.4: VIA OUTPUT 50

FIGURE 4.3.1: ARMA MODEL OUTPUT 51
FIGURE 4.3.2: ARMAX MODEL PERFORMANCE 52
FIGURE 4.3.3: VAR MODEL PERFORMANCE 55
FIGURE 4.3.4: NARX MODEL ARCHITECTURE 58
FIGURE 4.3.5: NARX PERFORMANCE VALIDATION 58
FIGURE 4.3.6 : TIME SERIES RESPONSE 59
FIGURE 4.3.7: REGRESSION OUTPUTS 59
FIGURE 4.3.8: HYPOTHESIZED MODEL 62
FIGURE 4.3.9: GLOBAL STOCK INDICES PERFORMANCE 63
FIGURE 4.3.10: STOCK PRICE SEARCH POPULARITY TREND 63

FIGURE 4.4.1: BCAD OUTPUT 69

FIGURE 4.4.2: BLOCKCHAIN TECHNOLOGY IMPACTS 70

FIGURE 4.4.3: TECHNICAL REQUIREMENTS HIERARCHY 71
FIGURE 4.4.4: STOCK PREDICTION FRAMEWORK........ 65

ii. List of Tables

Number	Description	Page number
2.2.1	Summary of literature survey	19
2.3.1	Variables and models	21
4.1.1	Comparison with BSE Sensex	44
4.1.2	NSE NIFTY Bank index statistics	45
4.1.3	Comparison of NIFTY Bank with global banking indices	45
4.2.1	VIA procedure results	50
4.3.1	ARMA Model Statistics	52
4.3.2	ARMAX model results	53
4.3.3	VAR system model	54
4.3.4	VAR model results	54
4.3.5	MLP model network information	56
4.3.6	MLP model summary	57
4.3.7	Predictive accuracy summary	81
4.3.8	Demographic distribution of countries	63
4.3.9	Global stock indices correlations	83
4.3.10	ANOVA	64
4.3.11	Bivariate Regression Coefficients	64
4.3.12	Model Summary	65
4.4.1	Top impact making 5 studies (source: Scopus)	85
4.4.2	Top impact making 5 studies (source: Web of Science)	87

iii. Nomenclature

Abbreviation	Description
AIC	Akaike Information Criterion
AMH	Adaptive Market Hypothesis
ANN	Artificial Neural Network
ARCH	Auto-Regressive Conditional Heteroskedasticity
ARIMA	Auto-Regressive Integrated Moving Average
ARMAX	Auto-Regressive Moving Average with exogenous inputs
ASEAN	Association of Southeast Asian Nations
AST	Adaptive Structuration Theory
BIC	Bayesian Information Criterion
BSE	Bombay Stock Exchange
BCAD	Bibliometric Coupling Analysis of Documents
CBR	Case Based Reasoning
CCAR	Co-citation Analysis of References
CNN	Convolutional Neural Network
DSS	Decision Support System
EGARCH	Exponential GARCH
EMH	Efficient Market Hypothesis
EPS	Earnings Per Share
ERP	Enterprise Resource Planning
GARCH	Generalized ARCH
GDP	Gross Domestic Product
GRETL	Gnu Regression, Econometrics and Time series Library
IBM	International Business Machines
ICT	Information and Communication Technology
IS	Information Systems
IT	Information Technology
LSTM	Long Short Term Memory

MCDM	Multiple Criteria Decision Making
MIS	Management Information System
MLP	Multi-Layer Perceptron
NARX	Non-linear Autoregressive Exogenous model
NASDAQ	National Association of Securities Dealers Automated Quotations
NSE	National Stock Exchange
NYSE	New York Stock Exchange
OLS	Ordinary Least Squares
P/B	Price-to-Book ratio
P/E	Price-to-Earnings ratio
PCA	Principal Component Analysis
PLS	Partial Least Squares
RMSE	Root Mean Squared Error
RWH	Random Walk Hypothesis
SEBI	Securities and Exchange Board of India
SEM	Structural Equation Modeling
SPSS	Statistical Package for Social Sciences
SVM	Support Vector Machines
TAM	Technology Acceptance Model
VAR	Vector Auto-regression
VIA	Variable Importance Analysis

INTRODUCTION

Chapter 1.
INTRODUCTION

Its common knowledge that the financial markets worldwide, especially stock exchanges of countries, exhibit random fluctuations due to diverse causal factors. While some of these are observable or inferred through empirical measurement, many variables have evaded closer scrutiny to researchers for a long time. Although the Amsterdam stock exchange existed around 1602 as the oldest stock market in recorded human history, empirical research relating to stock markets began only around the 1930s. “Stock exchanges” implies an open market only for “exchange” by buying or selling shares/securities entitling firm ownership for financial value.

A well-regarded economic principle underlying exchanges is the demand and supply theory. In any competitive market, the unit price of assets varies continuously unless the equilibrium point is reached. The majority of global markets are disrupted against reaching this equilibrium by seemingly unrelated events (Cowles, 1933; Agarwal *et al.* 2019). For example, during 1929, the great depression was signaled early from the crash of the London stock exchange, which was followed by the Wall Street crash in the USA. This phenomenon began as a causal chain of events probably triggered by economic reasons like speculative trading with other social and political factors but remained a highly debated issue.

Nevertheless, the role of stock markets as a significant barometer of economic activity and nations' overall financial well-being became evident. Before the turn of the 21st century, the notion of “physical” stock markets as a location for trading shares paved the way for electronic financial markets due to advances in Information Technology (Dewan and Mendelson, 1998; Loh and Ong, 1998; Greenwood and Jovanovic, 1999; Schinckus, 2018). One of the notable impacts was paradigm shift signaling that infrastructure investments in information technology and timely operations for the competitive advantage of firms. This relationship has arisen due to the pricing adjustment and delays occurring between information processing by stakeholders. The following subsection describes the background of the field broadly.

1.1. BACKGROUND

While many prior studies have shed light on specific markets prediction and models, a comprehensive analysis of sector-wise analysis and forecasting has been dearth. A few related research conducted in past decades is being summarized in this chapter. Investigations on possibilities of forecasting business with statistical information for stock markets existed among economic scientists (Brookmire, 1913). In the economics research field, Cowles (1933) quantified the financial service's recommendations of stock investment from 1929 to 1932 to statistically find that recommendations were worse than the average of common stocks annually. These results indicated that more profit-making recommendations were by chance rather than skill exhibited by analysts. With the development of digital computers and technologies, researchers started looking to apply them to practical problems. Studies in the computing field were also investigated like Hansen, 1956. His research mentioned the benefits of using a computer in stock market analysis from a technical perspective. A system was developed for the New York stock brokerage firm for the purpose. This study concluded that the analyst categories of tasks needed to be done; segregating technical and fundamental indicator variables and complexities involved are assumed in computation steps. The findings also mentioned challenges such as checking relevant theories and computation of all market-weighted averages.

From the business research field, an early study is by Jaffee *et al*. 1974. They opined that using a simulated investment strategy-based model continued to outperform a buy-and-hold strategy significantly. This strategy is buying stocks and not selling in the market, expecting an increase in their value by a trader, even after accounting for total transaction costs. Also, the perils of planning a strategy solely from stock prediction models were observed. Barrett and Wright, 1974, attributed the stock price increase to deterministic effects in the operations research area and suggested that the Brownian motion explains the price movements. Braun and Chandler, 1987 also devised a rule-based expert system that effectively generated profits in simulations from the decision sciences field.

Spann and Skiera, 2003 designed a virtual market for stock trading to understand forecasting and participant assessments about the management sciences field. Later studies adopted similar strategies of using simulation to understand complexity in real-world markets. Similarly, Patel *et al*. 2015 was the first major empirical review study from an Indian context from expert systems. In their research, methods like Artificial Neural Networks (ANN), Support vector machines, random forest, etc., were simulated for stock index predictions. Results noted that input parameters in the form of deterministic data greatly improved accuracy.

The studies that dwell into prediction models utilize variables that capture the market dynamics in both qualitative and quantitative factors. Correspondingly, these factors have been broadly classified into two types and adopted by industry analysts. Hence there have been majorly two branches of stock market analysis and or prediction; Technical analysis and Fundamental analysis. Technical analysis as the branch of the active investment plan and research gained after Levy (1966) gave foundational ideas involved in the process. Fundamental analysis has a genesis that goes back to the early 1980s, refined later by Spooner, 1984. During this period, investment management slowly began to evolve as an interdisciplinary field consisting of financial analysts, engineers, and industry professionals. A detailed discussion of these methods is provided in the second chapter.

1.2. **BANKING AND FINTECH**

In India, the services sector accounts for nearly 53.17% of the GDP as of 2017, and banking is a prominent component. Both the public and private firms listed in stock markets (BSE or NSE) raise a significant portion of capital for business through trading. As firms issue many shares collectively, they can be grouped in the stock index for sectors (Bonanno *et al*. 2000).

The current study exclusively focuses on BSE Bankex or NSE Nifty Bank (Balaji et al. 2018; Maji et al. 2017). Banking is influenced by technology changes and tends to adopt innovations much faster than other sectors due to competitive advantage and mitigate potential risks. The emerging paradigm has given rise to a field called financial technology (FinTech). A popular definition of FinTech is "*the new applications, processes, products, or business models in the financial services industry, composed of one or more and provided as an end-to-end process through the Internet and used to computerize insurance, trading, and risk management*." New FinTech companies and market activity are reconstituting the competitive landscape, changing the definition of a player in the financial services sector. Financial technology (FinTech) has appeared as a relatively new industry in India. It has companies that use technology to offer financial services. These companies operate in insurance, asset management, and payment, etc.

More importantly, recent research mentioned that a positive relationship exists between FinTech funding and the stock market performance of retail banks (Li *et al*. 2017). This area of research investigates new technology and innovation that aims to compete with traditional financial methods in the delivery of financial services. Smartphones for mobile banking or investing services are examples of technologies aiming to make financial services more accessible to the general public. As a result, FinTech is blurring lines between technology and traditional finance. It is a rapidly evolving segment of the financial services sector that enables start-ups and other market participants to disrupt the traditional financial services industry (Gabor and Brooks, 2017).

1.3. **ROLE OF ANALYTICS AND DSS**

The worldwide stock markets are highly complex, stochastic, and non-linear systems in varying magnitudes over time (Palsson *et al*. 2017). Under this assumption, predicting or even modeling with fair accuracy is challenging. Still, this is one of the highly debated topics in finance, economics, and computing (Junqué *et al*. 2013; (Chang and Ramachandran, 2017). The data mining tools such as Google trends and algorithms can extract predictive information about economic indicators from public internet trends (Choi and Varian, 2012). The benefits of a recommendation system practically effective enough for generating profits mean a lot, especially for investors, markets, companies, and other stakeholders who constitute this environment (Dimpfl and Jank, 2016).

As discussed earlier, the stock markets have a vast stake in economic systems around the world. More importantly, the fact is that significant role plays for banking & finance industry in crisis scenarios of global nature ex: regulation of recession during 2007 or vital task of risk mitigation in future events (Haldane and May, 2011). The additional information of public web data augments analysts' forecasts under volatility with higher accuracy and robustness for managerial decision support (Attigeri *et al*. 2015). Hence, business intelligence tools and applications can improve positive implications for the banking networks (Hu *et al*. 2009; Moro *et al*. 2015).

Apart from FinTech, the role of Information systems (IS) has been critical to organizations. However, its application through Information and Communication Technologies (ICT) faces a host of institutional and environmental challenges, leading to mixed outcomes (Venkatesh *et al*. 2016). Also, Decision Support Systems (DSS), a category of information systems application, are designed for specific tasks within this area like stock prediction, investment decision-making support etc. (Cho, 2010). In banking, the design of these information system performances significantly impacts overall economic and organizational performance (Davamanirajan *et al*. 2006). In the following subsection, the key motivations and scope of the research study are being described.

1.4. **MOTIVATION AND SCOPE**

After inferring from prior literature, many limitations are observed in models tested by researchers. At the same time, a considerable impact is in terms of performance and regulatory outcomes. While the theory is embedded as an explanation, specific applications assuming the banking and financial services industry as a case study for forecasting stock performance are dearth. Additionally, the financial research literature has shown a positive effect of stock markets and banking development on long-term economic growth (Levine and Zervos, 1998). Such a relation is consistent even if economic and political factors are controlled. Similarly, Cooper *et al*. 2003 opined that the fundamental variables could effectively predict banking stock returns.

It's found that quarterly changes in price-earnings and price-book ratios and asset calculation strategies can predict cross-sectional bank stocks (Mohapatra and Misra, 2019). Additionally, this study shows that quarterly data tends to get affected by sector-led focus, policy rate on a short-term basis, non-performing assets, etc. From these aspects, the design of predictive information systems has become vital for organizations that utilize analytics in decision-making, as seen in earlier literature (Kidd and Morgan, 1969; Shmueli and Koppius, 2011).

The earliest study on stock market prediction from a DSS perspective from India was by Rihani and Garg, 2006. Before that or even later, there has been scant literature on standard models or prediction models within the Indian context (Panda and Narasimhan, 2006; Mohapatra and Misra, 2019). Due to this, the current research study attempts to systematically review existing literature & models, find research gaps, and address them from the purview of information systems applications. As opined by Gomber *et al.* (2017), the future of FinTech solutions will be driven both by innovations on the technology level and by the industry reaction and regulators' assessment of the new developments. Customers will appreciate technological solutions that ease usage and reduce transaction costs.

Empirical research can utilize by addressing local concerns to proliferate FinTech (Gupta and Xia, 2018). Due to the stochastic nature of variables affecting fluctuations in the stock market data, deep learning has also emerged as key technology but still faces challenges (LeCun *et al*. 2015; Thakkar and Chaudhari, 2021b). The advantages of such models are powerful feature learning capability and robust results that can be applied for diverse tasks within the stock market or outside also. While many models were proposed in the literature, skepticism has arisen about accuracy and efficiency due to mixed results (Ding *et al*. 2015). Stock market prediction is considered a challenging and wicked problem because of no concrete or global solution, mainly resulting from the high stochastic nature of data.

Moreover, there is an apparent impact of the stock market on future economic growth (Pradhan *et al*. 2014). Several factors have been investigated, like micro-economic conditions, macro-economic variables, political scenario, investment friendliness, etc. Further, an index portfolio based on deep learning has shown good performance on tracking like the Hang Seng index (HSI) of Hong Kong (Ouyang *et al*. 2019). However, an industry analyst uses past historical market data, current conditions, and expert intuition to forecast estimates (Nofer, 2015).

Hiransha *et al*. 2018 used simulated models on data set both from NSE and the New York stock exchange. They suggested that CNN (Convolutional neural network) is more effective in both experiments than ARIMA. Singh and Srivastava, 2017 used deep learning methods on NASDAQ stock data to predict Google stock price. Another study by Balaji *et al*. 2018 used the data from the banking sector of BSE for prediction. It opined that for 4-step and 1-time step ahead stock directional movements, extreme learning machine (ELM) based models gave the best performance. Fawaz *et al*. 2019 reviewed time-series classification for data mining-based applications. Borovkova and Tsiamas, 2019 had ensemble methods LSTM (Long short-term memory) for high frequency. The focus on market analysis and others drawn from earlier innovations forayed to industry (Li *et al*. 2018; Levinson, 2005). Recent work utilizing tick data of 1 month period is shown to be efficient over 15-min test cycle for market forecasts (Selvamuthu, 2019).

While predictions are possible, making inferences on underlying causal mechanisms is tricky due to time-dependent random effects in diverse market sectors. As a modeling tool, the Partial least squares (PLS) method helps to overcome this by considering latent variables into this framework comprising stock market drivers and performance. The theoretical background and algorithm of the PLS algorithm were described in Wold (1954). According to Judea Pearl, an Association of Computing Machinery (ACM) Turing awardee, works of Wold on causal inference, graphical models, and observational studies were far ahead of their time (Pearl, 2019). The PLS method was adopted from many fields such as anthropology, bioinformatics, neuroscience, etc. However, one of the most popular fields is chemometrics. Path analysis was developed with roots from sociology by Duncan (1966), later applied to biology, psychology, etc. Structural Equation Modeling (SEM) evolved from path analysis to a standard statistical procedure. Similarly, empirical works on predictive information systems in management and systems discipline started relatively early (Kidd and Morgan, 1969).

French (1972) looked at multivariate behavioral aspects by conducting factor analysis on stock price changes for prediction purposes. A study by Kim and Park (1996) built intelligent decision rules for stock trading using such a method. Keller and Siegrist (2006) empirically demonstrated that the financial risk attitude and income were significant positive predictors for stock investments in Switzerland. One of the long-held view that investment decisions are based on returns or risks were challenged by Aspara and Tikkanen (2010). A study found that personal relevance and individual affective evaluation of brands are essential attributes affecting investment attraction in the Helsinki stock exchange. Another survey by Meenakshi and Lakshmi (2013) investigated the efficacy of a computational model for stock prediction in the Indian context combining neural networks and PLS that reported positive results. Utilizing data from the Spanish stock exchange, Pascual *et al*. 2014 proposed a model under the Theory of Planned Behavior (TPB) that explains 2/3rd of variance of intention and half of variance among attitude and control behaviors. Other works have focused on behavioral and social aspects like investor information sources (Khan *et al*. 2017), sentiments, financial knowledge, etc., among others.

With the proliferation of electronic mediated market transactions, news events (Guldiken *et al.* 2017; Verma *et al.* 2017), social media, the sheer scale and complexity involved increased exponentially. Shiller and Pound (1989) analyzed ten stocks listed in the Securities and Exchange Commission, 30 institutional investors, and 134 individual investors. They provided empirical results that questioned the tenets of the efficient market hypothesis inferred through survey data. For institutional investors, the rapid interest and price increase was influenced by individuals or publications. Individual investors were influenced more by friends/family, and decisions less affected by stockbrokers. East (1993) extended the study on investment decision-making under the purview of the theory of planned behavior perspective.

Similarly, Abreu and Mendes (2012) surveyed 1559 individual Portuguese investor's securities and found empirical evidence for a strong and positive relationship between investments in information and trading in financial assets. Also, based on the level of confidence exhibited by investors, the frequency of trading and advice soliciting can vary. As seen from the literature review, a comprehensive study must be required that examines theoretical dimensions of investment behaviors. Additionally, most previous works focused on models tested in broad markets or indices. While this is the case, experiments and analysis specified in the Indian scenario are missing; hence addressed in the current research. At the same time, quantitative variables are considered within the architectures of stock market prediction. So empirical research can augment the decision support models, assuming market behavior and forecasting for banking as a sector. Separate studies have investigated the usage of Artificial neural networks for prediction. These were positioned taking data of individual bank firms or using other financial ratios/economic variables. Of around 230 studies surveyed, only 18 explore prediction models tailored for this sector (~ 7.5%). More importantly, decision support systems for investment are not yet proposed from an Indian perspective. These have been implemented elsewhere, ex: Cho (2010) for China, Feuerriegel and Gordon, 2018 for German/European indices, Gottschlich and Hinz, 2014 for Germany, Leigh *et al*. 2002 for the NYSE (USA), Samaras *et al*. 2008 proposed bank stock trading for decision support model in Athens stock exchange.

1.5. **ORGANIZATION OF THESIS**

This thesis is being structured into a total of 5 chapters and 15 subsections; 1^{st} chapter is an introduction to the broad topic. Here the genesis of the topic is described in a detailed fashion. The motivation for research work is also mentioned.

The 2^{nd} chapter reviews the literature and the theoretical background of intelligent information systems and stock market analysis methodologies. Subsequently, the knowledge gaps are identified, and research objectives and questions are framed. Later, the design of the research is done, and the hypothesis is formulated.

The 3^{rd} chapter elucidates the research methodology covering models of study, the data sources, and software tools.

The 4^{th} chapter dwells on the detailed data analysis and interpretation. The simulation methods and models, performance details are compared and discussed.

Finally, the 5^{th} chapter gives a summary of significant findings, and the conclusions are derived. Also, limitations and future directions for research for academics and industry practice are acknowledged.

The Appendices lists two tables, references used for study, publications, and the researcher's resume.

STOCK MARKET MODELING- A CONCEPTUAL FRAMEWORK

Chapter 2.
STOCK MARKET MODELING: A CONCEPTUAL FRAMEWORK

This chapter outlines the previous studies that have been carried out on the stock market prediction problem. The researcher conducted an extensive and systematic survey on the articles and models proposed in the relevant literature. The timeline of analyzed works ranges from the year 1925 to 2021. The focus is given to primary literature appearing from top peer-reviewed journal articles, premier conference proceedings. At the same time, secondary and tertiary sources such as book chapters, industry reports/white papers, patents, etc., have also been utilized.

2.1. THEORETICAL PREMISE

In this subsection, the major financial theories, namely Efficient Market Hypothesis (EMH), Random walk theory or Random Walk Hypothesis (RWH), Adaptive market hypothesis (AMH), and Noisy market hypothesis, are defined and explained. Among the type of analysis, the prediction methods are also elaborated. Researchers and industry analysts perform stock market analysis and prediction using two methods; Fundamental and Technical analysis. Fundamental analysis is a method of analyzing stocks by calculating a stock's intrinsic value (Allen and Taylor, 1990; Samaras *et al.* 2008). Financial analysts using this are researching everything ranging from the general economy and industry dynamics to current financials or management (Figure 2.1.1).

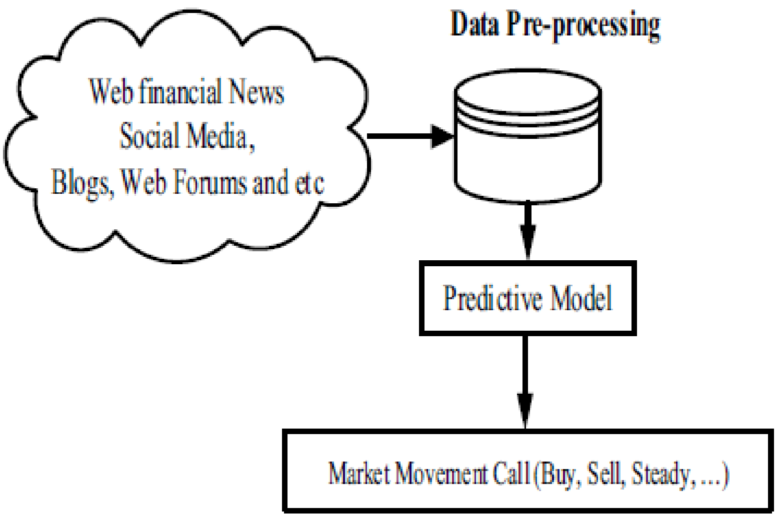


FIGURE 2.1.1: FUNDAMENTAL ANALYSIS (SOURCE: NTI ET AL. 2019)

As seen in Figure 2.1.1, the fundamental analysis utilizes various data sources such as web financial news, social media, blogs, web forums, etc., for market stakeholders. Whenever analysts perform estimates or market forecasts, this information is taken into account. At the first stage, the data collected is preprocessed. These include data cleaning, filtering, outlier detection and handling, and sparse data removal. Such kind of pre-processed and high-quality data is essential for analysis or prediction purposes. Subsequently, this data is fed into a predictive model that can estimate future market conditions, both short-term or long-term.

Big data research refers to a similar approach by collecting and analyzing data from diverse sources, including primary methods such as surveys, interviews, etc., and secondary sources like the public web, machine logs, sensor data, etc. (Thakkar and Chaudhari, 2021a; Nti *et al.* 2021). Technical analysis differs in principle that only inputs are the price and volume of the stock. It has evolved from ideas proposed in earlier literature (Hansen, 1956; Levy, 1966; Chenoweth *et al.* 1996; Lo *et al.* 2000; Neely *et al.* 2014). Here the central notion is that all established factors are factored to price, so other factors are not paid too much attention. Instead, the technical analysts ignore the intrinsic value and use stock charts to identify patterns trends (Figure 2.1.2).

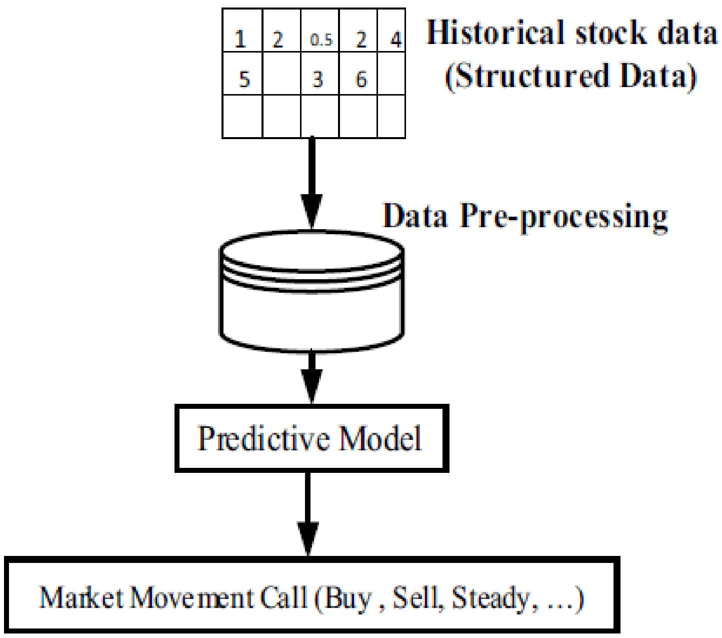


FIGURE 2.1.2: TECHNICAL ANALYSIS (SOURCE: NTI ET AL. 2019)

As shown in Figure 2.1.2, the technical analysis approach takes a conservative approach in the amount of data. Most importantly, historical data about the stock markets as the secondary source are preprocessed, including sparse value removal, outlier detection, normality checks, and filtering. A predictive model construction follows the data processing, and output indicates the market movements. In the case of classification models, only two outputs are generated, i.e., up/down. While estimation or regression models are used, quantitative value forecasting future market information is generated. These could be used for decision-making factors by the various stakeholders.

The efficient market hypothesis (EMH) proposed by Malkiel and Fama, 1970 opined that the market reflects all available information at any given time and is efficient. Adaptive market hypothesis: Lo, 2004 reconciles EMH with behavioral economics and proposes that attributes such as overconfidence, loss aversion, and other biases in investors are consistent with evolutionary models of adaptation to financial environment changes. As per the Noise market hypothesis defined by Siegel, 2006 the financial markets, including the stock exchange, do not hold the actual prices of securities.

Trading also happens from momentum traders, speculators, or institutions that participate in the market for other reasons than fundamental value. Similarly, the Random walk hypothesis (RWH) earlier and explained by Cootner (1967) is the second primary theoretical school of thought. It argues that movements in the financial market are entirely random and impossible to estimate or predict. Understanding these theoretical premises is vital since the prediction purposes used in the model need to factor in aspects like the economy, overall investment nature, and policies in force. Much of these inherently affect the choices that individual investors make for trading decisions. For example, in emerging markets like India, investors may have more social circles and share information about stocks much more than in developed economies. As a result, the volatility of certain popular stocks tends to be higher during more demand like IPO (Initial Public Offerings). Essentially, the fundamental and technical analysis is analogous to qualitative and quantitative research approaches, respectively.

2.2.REVIEW OF RELATED WORK

For getting comprehensive insights, the review of literature is divided broadly into four periods. First is the earliest period covering 1913 to 1965 (52 years) to highlight initial developments within the topic. The second comprises of period 1966 to 1989 (23 years). The third period is from 1990 to 2005 and 2006 to 2021 (15 years). The scientific literature published in top journals/conferences/books, conferences, and patents is analyzed.

The majority of these sources were listed on Web of Science (SCI/SSCI), Scopus, DOAJ and ranked as per Chartered ABS or ABDC ranking list. Such papers were extracted based on overall theme, citation impact and match to area

i. Early Years (1913 to 1965)

Brookmire, 1913 investigated the issues inherent in forecasting for businesses. The work analyzed the economic and financial data to offer suggestions and the impact of explaining phenomena of prediction.

Schultz (1925) reported about the American Statistical Society meeting wherein security prices forecasting was the topic. It was opined that the stock market is a creation and resultant of economic forces. A barometer, one of the early forecasting methods by chart reading, was found to be inaccurate, termed as ‘diagnoses’. The report also documented various speakers and agencies that formed the industry.

Cowles (1933) explored the data from 16 financial services, about 7500 endorsements of individual stocks for investment from January 1, 1928, to July 1, 1932. Results compiled an average record worse than average common stock by 1.43% annually. Statistical tests were unsuccessful in proving that they exhibited professional skill and indicating this was perhaps the result of pure chance.

Von Szeliski (1956) attempted to construct a prediction instrument on 90 days stock from Standard and Poor’s, with the aim being five-week trend averages assuming movements in the same direction. The data used was from October 1932-January 1933.

ii. Secondary Stage (1966 to 1989)

Krolak *et al*. 1969, in one of the earliest empirical studies on stock market investment studies, used data from July to September 1968 on 20 stocks portfolio. The simulation used inputs profit/loss of transactions, number of shares, the average price per share and commissions/trade, etc. The system forecast results were generated as reports for the decision-making in the future.

Felsen (1975), noted the high complexity in the process of price generation, and designed an investment decision support system. Also, learning algorithm for stock trading was proposed. Based on computational approaches using data sets of the 1970 New York composite index, the study found that investment analysis involves processing complex information patterns. Because of this fact, investment policy must reflect the number of social, psychological, technological, and other factors involved.

White (1988) used sample data by taking *r*, the one-day rate of return to hold IBM common stock on day *t*. From 5000 days of returns data, a sample of 1000 days for training purposes taken with samples of 500 days before and after the training period. Datasets were from 1974 through 1978. Suggestions aimed to include more inputs into the model, i.e., volume, extra stock prices and volume, prominent indicators, macroeconomic data, etc.

iii. Intermediate Period (1990 to 2005)

Refenes *et al*. 1994, under the context of arbitrage pricing theory, showed how sensitivity analysis overcomes issues of the black-box nature of neural network predictions. They identified performance factors and model parameters and explained multifactor model comparisons indicating better modeling performance of ANN systems.

Levin (1996) constructed a multilayer feed-forward neural network model to predict excess returns in a stock index portfolio. It used both fundamental and technical indicators apart from using performance measures of Sharpe ratio using Treasury bill data.

A benchmarking analysis was done with a generalized least squares model. Results indicated that such models could extract insights on market factors and expected returns even if the low signal-to-noise ratio of data affected inputs into the model. A similar study confirmed complexities within stock market time series (LeBaron *et al*. 1999)

Lo *et al*. 2000 explored technical analysis smoothing techniques such as nonparametric kernel regression and inferred a method. Using it helps to recognize predictabilities in the time series of stock prices by extracting non-linear patterns from noisy data. In their study, daily returns of individual NYSE and NASDAQ stocks from 1962 to 1996 were used along with bootstrapping and Monte Carlo simulation procedures. They suggested practical use of technical analysis can add value to the investment process.

Leung *et al*. 2000 while denoting a dearth of research on stock index movement, compares time series models based on performance accuracy and investment returns. A detailed examination was done on linear discriminant analysis, logit, probit, and ANN models. The level estimations like vector autoregression with Kalman filter, exponential smoothing, and multilayer feed-forward ANN were compared. Results showed that classification models are better to level estimation for stock trading.

Walczak (2001), perhaps in the first empirical study, highlighted the importance of management information systems (MIS) on stock predictive models. Also, counter-intuitive results were produced. Usage of a higher amount of training data does not guarantee increased predictive accuracy to estimate foreign currency exchange prices.

Kim (2003) implemented Support vector machines (SVM) based model that predicts the stock price index. The output of this architecture was compared with the backpropagation (BP) based model Case-based reasoning (CBR). The inferences revealed that though SVM outperforms BP and CBR identifying specific parameters was essential for better accuracy.

Lam (2004) analyzed an information system perspective that experimented with a nonparametric model such as ANN incorporating technical and fundamental factors. The input attributes for the predictive model were 16 financial statement variables and 11 macroeconomic variables.

Data of 364 S&P companies for the period of 1985 to 1999 was used. Results showed the capability of neural networks to considerably outperform the benchmark based on a highly diversified investment strategy. A combination of financial statements and macroeconomic variables based predictions doesn't guarantee significantly higher profit returns. Such integration better suits during an economic recession.

iv. Recent Advancements (2006 to 2021)

Dutta *et al*. 2006 used the closing prices in the weekly BSE SENSEX closing prices for 250 trading weeks starting from January 1997 up to 2003 to train the ANNs model. Findings show the error rate increases gradually during the period of validation. Hence, an appropriate approach may be to retrain the network periodically.

Samaras *et al*. 2008 had designed a Multiple Criteria Decision Model (MCDM), deploying the fundamental analysis exclusively for application in the Athens stock exchange. Their study developed a DSS specific in the banking sector for investors.

Atsalakis and Valavanis, 2009) covered more than 100 studies that use ANN-based modeling for stock index prediction in an extensive review. There was no single study from the Indian context. However, few related works were in the literature that focused on volatility estimation using econometric or decision models. Hence it indicated a vast gap in theory and practice for designing and performance validation for this task.

Cho (2010) developed a prototype architecture for the Multilevel and Interactive Stock Market Information System (MISMIS). Different data modeling methods like ARIMA and ANN were studied as a technique for forecasting stock and helping investment decisions with data from Hang Seng Index (HSI).

Guresen *et al.* (2011), in a comparative study, investigated Multilayer Perceptron (MLP), a dynamic artificial neural network (DAN2). The hybrid ANN with GARCH of daily foreign exchange rates of NASDAQ from October 7, 2008, to June 26, 2009, was used. Findings showed classical ANN model MLP outperforms DAN2 and GARCH-MLP only with a little marginal difference. Additionally, the GARCH model inputs had been affected by noise within data.

Koutroumanidis *et al*. 2011 study applied ANN's to estimate the stock prices of Alpha bank. They used data of daily closing prices from January 2004 to November 2005. Findings suggested that the non-stationarity of time series can be overcome using the bootstrap method and does not limit the confidence interval in forecasts. Comparisons using the hybrid approach, Kalman filters proved efficiency.

Nardo *et al.* 2016 looking into the impact of online financial news, used various methods and found mixed results. They opined that even though web activity does have predictive power on financial movement, gains rarely exceeded 5%. Baluch and Jackowska, 2018 experimented using fundamental and technical fractal analysis for stock market indices. Comparison of three hybrid models with results of ANN methods showed that the proposed approach gave better performance.

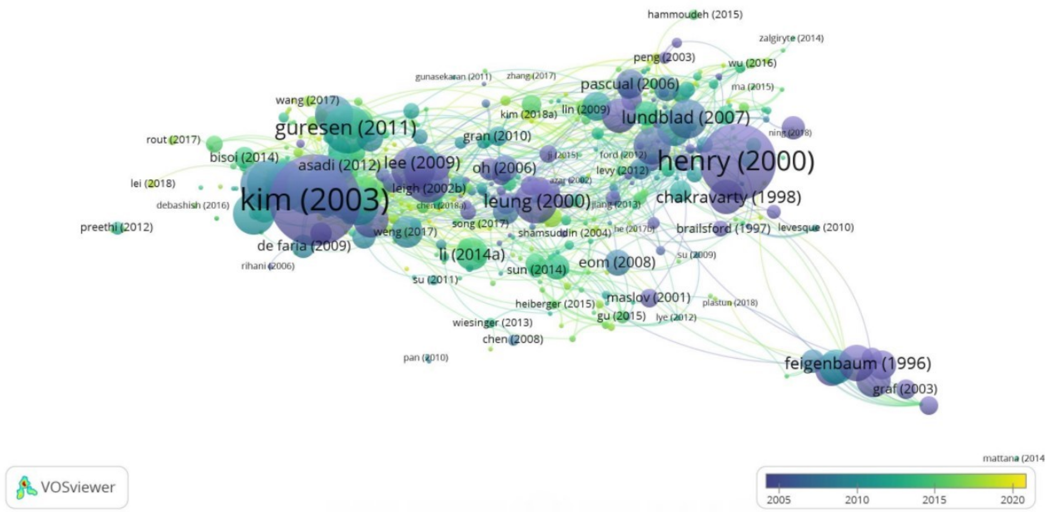


FIGURE 2.2.1: BIBLIOMETRIC COUPLING ANALYSIS (SOURCE: VOSVIEWER)

A detailed bibliometric analysis is carried out to understand better the existing knowledge base gaps in predictive models for stock prediction. For this, the bibliometric coupling analysis of documents (BCAD) method uncovers the current trend in field literature while visualizing impact and interlinks among individual studies (Figure 2.2.1). The bigger circles in the diagram indicate highly cited studies, and color-coding by software shows a timeline of published papers.

The high-impact study is by Kim (2003). The blue color dots indicate older research articles and yellow shows more recent ones. The Scopus database and Web of Science are being used to extract the relevant publications. The keyword used to retrieve results is "*Stock market index prediction*." From the results obtained, the document type is selected option "Journal" to filter such articles. Under this method, exploration of the corpus of literature included 545 articles, grouped in 15 clusters using association strength of similarity and total no. of reference links 16881. Interestingly, Atsalakis and Valvanis (2009), a major review study on current research, could not include a single Indian context study published before 2009. The analysis reveals that apart from technical implementations of prediction, the broader framework has been missing that encompasses organizational contexts and information system designs. Below table recap the findings.

Table 2.2.1: Summary of findings from prior literature

Prediction type	**Emerging markets**	**Developed markets**	**Limitations and Inferences**
Statistical/ Econometric	Birau *et al*. 2015, Malik *et al*. 2017	Chen *et al*. 2019	Significant forecasts of stock indices measured.
ANN models	Balaji *et al*. 2018;	Koutroumanidis *et al*. 2011	Only for individual bank firms and index omitted.
Hybrid	Dincer, 2015	Oztekin *et al*. 2016; Tan *et al*. 2007	Expert validation is hard.
Big data	Attigeri *et al*. 2015	Li *et al*. 2021	Needs rigor in its scope.

2.3. LITERATURE MAP

A literature map is drawn that helps to visualize different types of predictive models, methods used, and significant indicative works for practice (Figure 2.3.1).

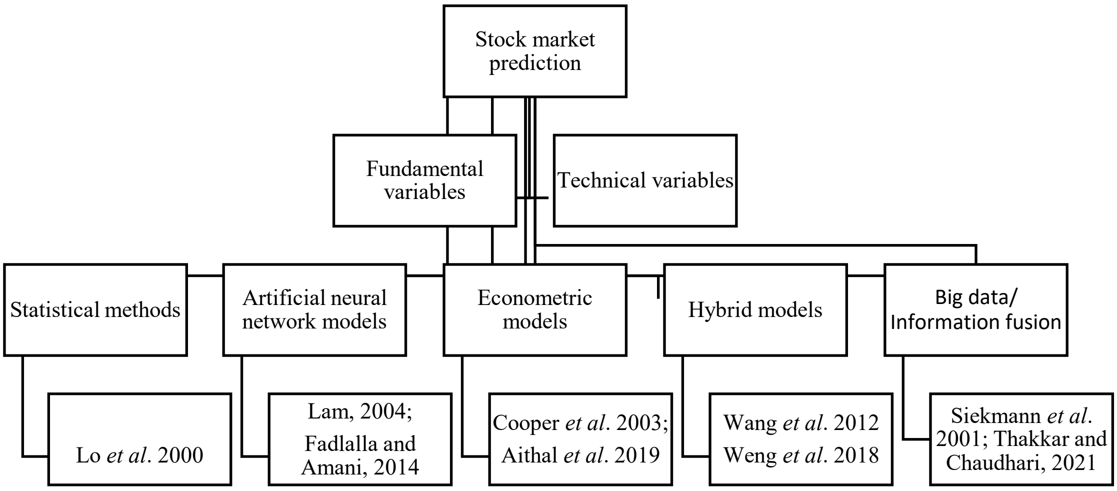


FIGURE 2.3.1: STOCK MARKET PREDICTION MODELS (SOURCE: LITERATURE SURVEY)

The figure only classifies indicative studies and not exhaustive listing characterizing various methodologies. Primary classification in its scope has been confined to purely statistical methods, artificial neural network predictions, and econometric methods. More recently, hybrid models and big data-based approaches have gained importance to overcome limitations in earlier techniques.

Table 2.3.1 lists the major variables explored in the literature to address stock prediction modeling. The factors assumed in forecasting, major works, and the models developed in prior research are categorized accordingly.

TABLE 2.3.1: VARIABLES AND MODELS

Variables included in the study	**Reference sources**	**Models tested in current research with supporting literature**
Price-to-earnings (P/E) ratio	Hickman and Petry, 1990; Zorn *et al*. 2014;	1. Automatic Linear Modeling (ALM) [Steyn *et al*. 2020] 2. Ordinary Least Squares (OLS) [Olson and Mossman, 2003] 3. Autoregressive Moving Average with exogenous inputs (ARMAX) [Ariyo *et al*. 2014; Babu and Reddy, 2014; Hiransha *et al*. 2018] 4. Vector Auto-Regressive model (VAR) [Khansa and Liginlal, 2011; Ibrahim *et al*. 2020] 5. Multilayer perceptron neural network model (MLP) [Sujatha and Sundaram, 2010a; Chen *et al*. 2019c] 6. Radial basis function network model (RBF) [Rout *et al*. 2012; Sheelapriya and Murugesan, 2017] 7. Non-linear autoregressive exogenous model (NARX) [Alkhoshi and Belkasim, 2018, Gandhmal and Kumar, 2021]
Price-to-book ratio (P/B)	Charumathi and Suraj, 2014; Mohapatra and Misra, 2019	
Dividend yield	Goyal and Welch, 2003; Baker and Belgorodskiy, 2007; Wu and Hu, 2012	
Stock index history (Open, Close, High, Low)	Leigh *et al*. 2002; Neely *et al*. 2014; Fadlalla and Amani, 2014	
Beta (Market risk)	Neely *et al*. 2014; Birau *et al*. 2015; Patra and Padhi, 2015	
Stock price search interest	Choi and Varian, 2012; Preis *et al*. 2013; Dimpfl and Jank, 2016; Perlin *et al*. 2017; Agarwal *et al*. 2017	

Source: Authors' review

2.4.RESEARCH GAPS

Based on the review of the literature, the following research gaps are identified. Such knowledge gaps pertain to theoretical knowledge in the academic field and practice for the industry. Correspondingly, research objectives are framed based on the gaps.

1. The majority of the studies have focused on choosing either purely statistical, artificial neural networks or econometric-based models. The hybrid modeling techniques have been experimented with in index prediction only around 2000 or later (Armano *et al*. 2005). In the case of Germany, Siekmann *et al*. 2001 looked into the possibilities of integrating various data sources for stock index prediction purposes. Wang *et al*. 2012 in their study, experimented hybrid model of stock index forecast. De Oliveira *et al*. 2013 tried to forecast the stock prediction performance with a firm-specific model of Petrobras. A study by Arasu *et al*. 2014 performed data mining methods to predict the stock market index in the stock exchange of Sri Lanka. Chen *et al*. 2019c designed a causal time series model using MLP to predict the Taiwan stock index. Likewise, Fadlalla and Amani, 2014 implemented a prediction model to estimate one day ahead of Qatar's closing price utilizing technical indicators and ANN. Kara *et al*. 2011 also investigated the stock index prediction in the Istanbul stock exchange using ANN and Support vector machine-based model.

   From all surveyed works, only 8 use Indian stock market data for the prediction model. That comprise 3.5% of total studies referenced for research. Due to these aspects, the banking sector index is chosen to support system support that also positions framework in the financial industry. Similarly, studies have explored psychology and organizational processes that affect predictive stock market-based decisions among individuals (Andreassen, 1987; Andreassen, 1988). Such findings hint at the process complexity involved within investors, stakeholders and human behavior. Later works have explored the phenomena on public mood and investor personality characteristics for stock market predictions (Chen *et al*. 2019a; Chen *et al*. 2019b; Earley *et al*. 1989; Jin *et al*. 2018)

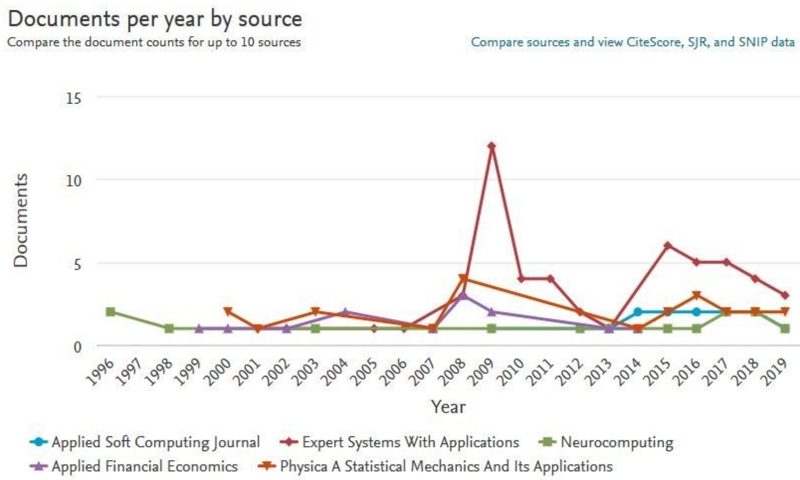


FIGURE 2.4.1: TIMELINE OF PAST RESEARCH (SOURCE: SCOPUS)

Figure 2.4.1 uses Scopus data to plot the trend of research published in journals. In 2009 high productivity was observed while it has lowered in recent years. It may also be attributed to the scholarly attention and industry requirement for better market prediction frameworks resulting from the global recession in 2008 (Weng *et al*.2018). Few prior studies address macroeconomic indicators that cause stock market indices to close. Still, hybrid-based models specific to prediction in India are very few.

2. Because of the nature of prior scientific literature, the statistical models used on forecasting stock market outcomes in the western economies need to be empirically compared for performance and accuracy in the emerging markets (Patel *et al*. 2015, Henrique *et al*. 2019; Moghaddam *et al*. 2016).

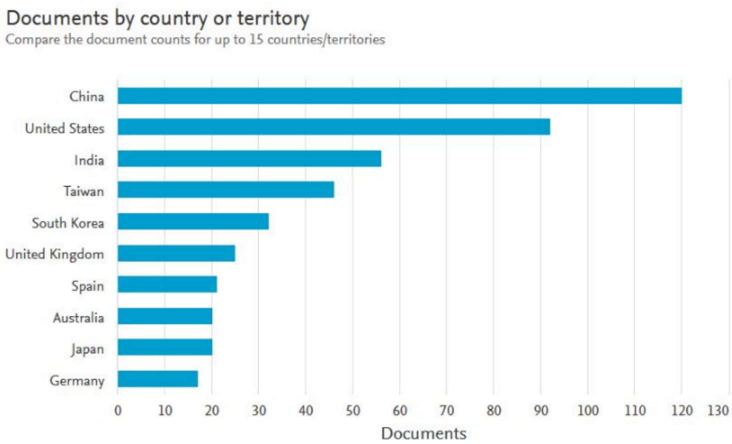


FIGURE 2.4.2: GLOBAL RESEARCH OUTPUT (SOURCE: SCOPUS)

3. Countries with a major contribution to the research field indicate that India is in the 3rd position after the USA and China (Figure 2.4.2). The artificial neural network model's applicability for modeling firms such as banking or specific stock indices reflecting this industry is rarely explored.

   In India, the banking part of the services sector accounts for more than half of the National Gross-domestic product (GDP) (Koutroumanidis *et al*. 2011; Charumathi and Suraj, 2014; Balaji *et al*. 2018; Pointer *et al*. 2020; Saud and Shakya, 2020). Thus, factors that affect the stock market index specific to Indian indices or financial firms are absent in the literature. Those available to date except for early studies are renewed interest in the topic (Rihani and Garg, 2006; Panda and Narasimhan, 2006; Dutta *et al*. 2006; Thenmozhi, 2006). From analyzed literature, 22 works use Indian stock market-related data for prediction. Within it, 3.5% are of works focused on banking. Hence, the banking sector index is chosen for research.
4. Finally, the earlier research has majorly focused on quantitative methods on secondary financial data. In contrast, qualitative factors such as investment environment and behavioral aspects are also vital for predicting future market fluctuations (Preis *et al*. 2013; Perlin *et al*. 2017). Such approaches gathering investor interest gathered through search behaviors etc., were experimented at developed markets though emerging countries lack such studies. Hence proposed research includes technical and fundamental indicators with qualitative data that impact the stock market indices for banking in India.

2.5. RESEARCH QUESTIONS

Four primary research questions are being addressed in the current study:

1. What are the major factors that affect the stock value index of banking firms in India?
2. Which statistical models are efficient in predicting the stock price index?
3. What type of neural networks can be used in the financial modeling of these firms?
4. How does the artificial neural network model perform comparing to conventional models?

2.6. RESEARCH OBJECTIVES

Few research objectives were formulated based on gaps. These are listed below:

1. To measure the importance of fundamental and technical indicators that affect the stock index of banking firms in the Indian services sector.
2. To evaluate statistical techniques in predicting the stock index price of BSE and NSE.
3. To compare the different artificial neural networks in stock index prediction based on forecast accuracy.
4. To create an optimum artificial neural network-based prediction model for forecasting the stock price index in this sector.

2.7. RESEARCH HYPOTHESES

To answer the questions, a few research hypothesis had been formulated:

- H1 (Null hypothesis): There is no significant difference in the estimation of stock index close of banking sectors using fundamental and technical indicators.
- H1 (Alternate hypothesis): There is a significant difference in the estimation of stock index close of banking sector using fundamental and technical indicators
- H2 (Null hypothesis): There is no significant difference in prediction accuracy of stock index close of statistical model and ARMA based model.
- H2 (Alternate hypothesis): There is a significant difference in prediction accuracy of stock index close of statistical and ARMA based model.
- H3 (Null hypothesis): There is no statistically significant relation of the global stock price search interest on the Indian stock market index close.
- H3 (Alternate hypothesis): There is a statistically significant relation of the global stock price search interest on the Indian stock market index close.

In the following chapter, the methodology adopted for the current research study is explained in detail.

R E S E A R C H

M E T H O D O L O G Y

Chapter 3.

RESEARCH METHODOLOGY

This chapter explains the overall methodology of research adopted for the study. Here, the research approach, a conceptual framework, is described in detail. Additionally, the nature and sources of data (both primary and secondary), sampling design, and software tools used for analysis are outlined. Chottiner (1972) described a stock market as a system, suggesting internal methods to gain insights into its working.

3.1. VARIABLES

The variables used in the research is being described in detail in this section.

i. Independent variables

3.1.1. **Price-to-earnings ratio (P/E)**

As one of the significant market valuation variables, the P/E ratio has much significance among stakeholders and industry analysts. Zorn *et al*. 2009 empirically showed how the extent of the changes in P/E has predictive power. Even though the variable has been utilized for general estimation in the extant literature, there is a shortage of studies that explore it for stock prediction tasks. Generally, it is defined as an estimate of a company's current price concerning its per-share earnings, defined by equation 3.1.1.

$$Price\ earnings\ ratio = \frac{Market\ value\ per\ share}{Earnings\ per\ share\ (EPS)} \qquad \text{(Eq. 3.1.1)}$$

Where, Market value per share is the market price of concerned share and EPS ratio.

Here EPS is further evaluated as in equation 3.1.2:

$$EPS = \frac{Net\ income - Peferred\ dividends}{End-of-period\ common\ shares\ outstanding} \qquad \text{(Eq. 3.1.2)}$$

The EPS as a parameter is used in earlier research works for valuation and prediction (Charumathi and Suraj, 2015; Jadhav *et al*. 2015).

As per prior financial research literature, evidence suggests that a higher P/E ratio necessarily may not attract investment since this imply stock as overvalued in markets.

3.1.2. **Price-to-book ratio (P/B)**

In predictive estimation, analysts use book value, as seen in many earlier studies. It's a ratio used to compare a stock's market value to its book value. It is calculated by dividing the current closing price of the stock by the latest quarter's book value per share. Also known as the "price-equity ratio". It is calculated as given in equation 3.1.3:

$$Price\ book\ ratio = \frac{Market\ price\ per\ share}{Book\ value\ per\ share} \qquad \text{(Eq. 3.1.3)}$$

Here the denominator is calculated as by equation 3.1.4:

$$Book\ value\ per\ share = \frac{Total\ Assets - Total\ Liabilities}{Number\ of\ shares\ outstanding} \qquad \text{(Eq. 3.1.4)}$$

This indicator was used in market prediction models in multiple works (Charumathi and Suraj, 2014). The three-factor model devised by Fama and French also included price-to-book value. Research has shown lower P/B ratio stocks could perform higher than high P/B valued stocks. Also, the variable reflects the valuation of industries with high intangible assets, such as the banking industry. In comparison, their liabilities and assets are weighed on market values.

3.1.3. **Dividend yield**

This variable is defined as the financial ratio that indicates how much a company pays out in dividends each year relative to its share price. Goyal and Welch, 2003, indicated that the rising persistence of the dividend-price ratio is primarily responsible for the dividend ratio's inability to forecast equity. But dividend rates also have strong long-term forecasts of 5-10 years. Similar works have explored the usage of dividend to price factor in prediction (Favero *et al*. 2011; Wu *et al*. 2012)

Dividend yield (D.Y) is represented as a percentage. It can be calculated as dividing the dollar/rupee value of dividends paid per share of stock held by the dollar with the rupee value of one share of stock. The formula can be shown as in equation 3.1.5:

$$Dividend\ yield = \frac{Annual\ dividents\ per\ share}{Price\ per\ share} \qquad \text{(Eq. 3.1.5)}$$

The efficacy of the variable as the predictive indicator has been shown in earlier literature (Hickman and Petry, 1990; Goyal and Welch, 2003; Baker and Belgorodskiy, 2007) with mixed results.

3.1.4. **Beta**

The market outcomes are expected to be influenced by risk factors. The market risk or volatility is denoted with the symbol beta (β). So, beta measures the volatility or systematic risk of a security or a portfolio compared to the entire market. It can be calculated using regression analysis.

A formula also is applied by analysts as in equation 3.1.6:

$$Daily\ beta = \sqrt{(\sum(P_{av-}\ P_i)^2}/\mathrm{n}) \qquad \text{(Eq. 3.1.6)}$$

Here P_{av} is the average price, and P_i is the price on an i^{th} day. It is calculated by first dividing the security's standard deviation of returns by the benchmark's standard deviation of returns. The resulting value is multiplied by the correlation of the security's returns and the benchmark's returns. A beta value b equal to 1 indicates that the security's price moves with the market. A beta $b< 1$, means that the security is theoretically less volatile than the market. A beta of $b> 1$ indicates that the security's price is theoretically more volatile than the market (Birau *et al*. 2015)

3.1.5. **Stock index history**

Almost the whole of the technical analysis heavily depends on this secondary data. The stock index price history consists of four data attributes Open, High, Low and Close. These data values for a particular index can be calculated daily, monthly, or yearly and used for predictions (Patel *et al*. 2015). Similarly, most surveyed studies have indicated a moderate to significant impact of the past performance of the future stock market.

3.1.6. **Stock price search interest**

Stock price search interest as a predictor has been studied by Dimpfl and Jank, 2016 and suggested that the benefits of the recommendation system practically effective for generating profits have high value, especially for investors, markets, companies and stakeholders constituting this environment. Preis *et al*. 2013 showed that Google search queries have important information containing investor's trading behavior on the global markets in an empirical study.

Later, this result spurred other works like Perlin *et al*. 2017 that proved that such information could be successfully utilized to forecast or model financial markets in the world. The data from Google trends is collected for using price search interest as a qualitative predictor. Further details about the data collection algorithm are explained in the next chapter.

ii. Dependent variable

3.1.7. **Closing index**

The closing index is the price of the last transaction of a particular stock index completed during a day's trading session on an exchange. Its computed daily, monthly or annual frequency basis (Fadlalla and Amani, 2014; Nayak *et al*. 2015; Gao *et al*. 2016; Weng *et al*. 2018). Hence, this study initially identifies major fundamental and technical factors among these chosen variables. Most of the prior literature has focused on generalized stock market prediction. But notably, the market closing is an essential aspect for stakeholders and analysts alike that have been overlooked.

The recent works investigated closing market impact like Weng *et al*. 2018; Nayak *et al*. 2018. Also, including more other variables can affect the model degree of freedom and exhibit overfitting issues reducing prediction accuracy/robustness. In Figure 3.2.2 in the next section, the overall strategy of research is depicted using a flowchart.

3.2. **RESEARCH MODELS**

In this section, the various predictive models being experimented with are described. Subsequently, the input data sets, methodology, architecture and results are interpreted. An artificial neural network (ANN) is a massively parallel distributed processor with a natural propensity for storing experiential knowledge.

Fundamental research into artificial neural networks dates back to 1943, with the seminal work by McCulloch and Pitts published in Bulletin of mathematical biophysics. In 1958, the concept of perceptron was developed by Rosenblatt published in Psychological Review journal. Following the work by Minsky and Papert in 1969, the research efforts were reduced mainly due to the discovery that perceptron's could not process data in exclusive-OR logic, as well as the limitations on computational power. However, progresses in transistors and digital technology again fueled more work in the field. Algorithm-based advancements such as backpropagation were developed by Rumelhart *et al*. in 1986. Hence this study also extends these for prediction purposes.

3.2.1. Multilayer perceptron models

Many neural network models exist in the scientific literature, having diverse applications. MLP (Multilayer perceptron) based neural network model implements an internal algorithm, i.e., backpropagation. Below is the outline for one input and one output network model. These can overcome limitations in statistical models to the best possible extent. An example of a neural network system is given in Figure 3.2.1.

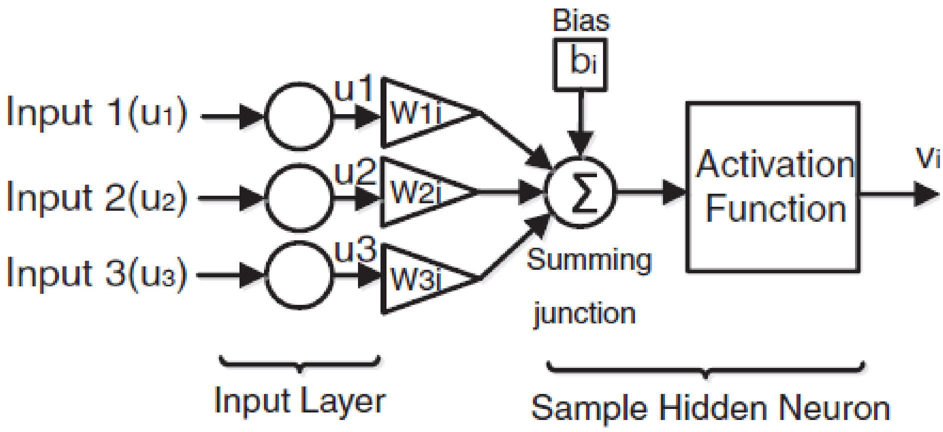


FIGURE 3.2.1: A GENERIC NEURAL NETWORK (SOURCE: ZEINALIZADEH ET AL. 2015)

3.2.2. Backpropagation (BP) algorithm:

Assuming a network with a single real input *x* and network function *F*. The derivative *F'(x)* is computed in two phases: Feed-forward: the input *x* is given input to the network. Now, primitive functions at each node and their derivatives are evaluated at each node. The derivatives are saved. The constant value one is fed into the output unit, and the network is run backward. Incoming information to a node is added. The result is multiplied by the value stored in the left part of the unit. The result is transmitted to the left of the unit. The result collected at the input unit is a derivative of network function with respect to value *x*.

As seen in Figure 3.2.2 given in the later section, the researcher adopts the process in designing the study. Once the research questions are framed, an appropriate sampling method, i.e., if a time-series (Start date-end date) is used for data, is chosen. The sampling process, in this case, has to cover the entire period of existence of particular stock indices identified. Once the information is collected, a visualization method needs to be used. These range from charts and plots to tables and figures. Before going ahead with the modeling part, the study needs to identify variables used in building the model, which is done at the fourth level depicted in the center box. After the variable selection process is done, this research study performs multiple simulations, i.e., neural networks and statistical model implementations.

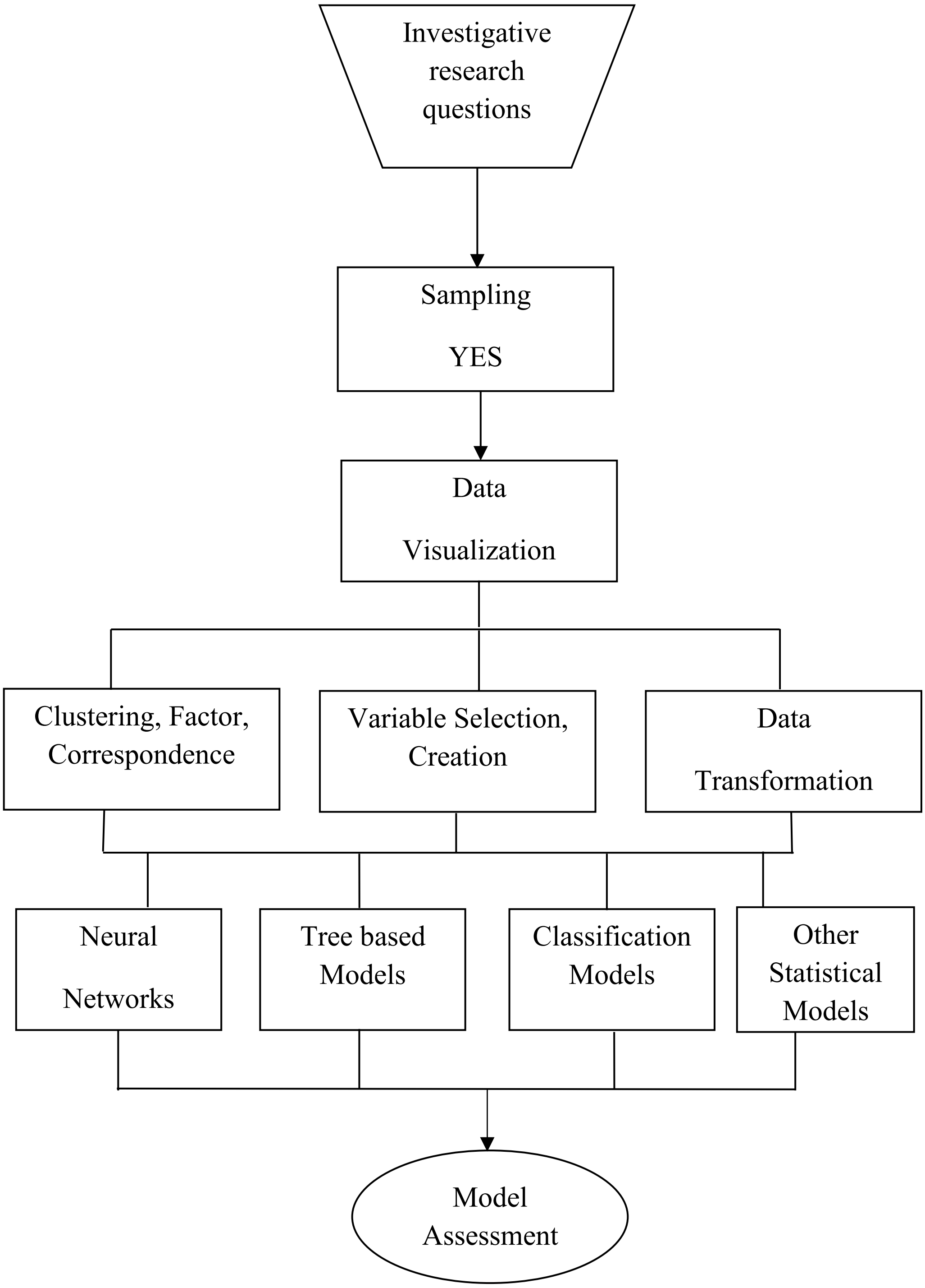


FIGURE 3.2.2: RESEARCH DESIGN (SOURCE: COOPER AND SCHINDLER, 2015)

3.3. SAMPLING DESIGN

The stock returns of banking firms listed in both BSE and NSE indices are collected and analyzed. Both public sector and private sector enterprises are included in the sample. Up to 14 years of data (2005-2019) on yearly, monthly, and daily frequency are analyzed. Artificial Neural Networks needs data specifically for training, testing and cross-validation of the results, which is followed here (Walzack, 2001). Out-of-sample forecasting is tested by default within IBM SPSS, MATLAB during simulation exercises. It ensures performance & robustly handles outliers.

3.4. DATA SOURCES AND TOOLS

The secondary data, i.e., stock prices of major banking firms publicly traded (both BSE and NSE based on net worth, inclusive contribution to economic growth, and dominance over the Indian equity market. The data during 2005-2018 on fundamental and technical analysis variables (P/E, P/B, Dividend yield, and Stock index historical data on Open, High, Low and Close attributes) are extracted from official stock exchange sources www.bseindia.com and www.nseindia.com/ for BSE and NSE respectively. The data in daily, monthly or yearly frequency is available for download in Excel file format.

An official audit verifies this data by the Securities and Exchange Board of India (SEBI) & CRISIL (Credit Rating Information Services of India Limited) database. Additional data obtained from Google Trends as stock price search behavior is used as a qualitative factor to optimize the predictive model. Interestingly, Patra and Padhi, 2015 showed that BSE Bankex data exhibits long-memory properties and effects of asymmetry. Due to this presence of autocorrelation exists in closing share prices of banks. Sudhakaran and Balasubramanian, 2016 suggested using data from 2005-2015 that FDI (Foreign Direct Investment) does not improve the stock index's performance, even having a negative effect. In contrast, the foreign exchange rate makes a 1% increase on indices.

3.4.1. Simulation and data analysis

MATLAB R2010a is a version used for statistical treatment, analysis, visualization of data. MATLAB has an inbuilt artificial neural network simulation feature following a static method and runs experimental simulations with trial data sets. The IBM SPSS version 21 also has neural network simulation, which is used in the present study. GRETL version 2016d is used for the majority of econometric analysis and modeling functions. Additionally, Microsoft Excel 2007 is being used for data tabulation, cleaning and pre-processing since the secondary data are in .xls/.xlsx format. For conducting the bibliometric analysis, the open-source tool VOSviewer is utilized.

3.4.2. Model implementations

Few reasons were evident to necessitate experiments with a few prediction models. ARMA model in earlier works exhibited good approximation fitness in the forecast of stock returns and indices. VAR is a proven method for stochastic modeling of phenomena that can include economic, financial and other factors. RBF was chosen as it gave a positive result in stock forecasting in at least two empirical studies for Indian stock market data. NARX model experiments are done as the model has a good convergence rate and supports the backpropagation algorithm.

3.4.2.1. ARMA model

ARMA is a time-series model for the prediction that classifies the particular random effect model with constant. Like this approach, Srivastava (2017) used ARIMA (0, 0, 0) as a reflexive model to investigate the extent of social investments comparing results to ANN. Similar model performance and accuracy designed for the stock prediction task have been modeled. Following the results of Hiransha *et al*. 2018, the ARMA with multiple indicators has increased predictive accuracy. They adopted a deep-learning NIFTY prediction model in 3 sectors and found ARIMA as a univariate model. It also captures dynamics to a lesser extent in multivariate time-series predictions. The notation ARMA (*p*, *d*, *q*) refers to a model with *p* autoregressive, *q* moving-average terms, *d* differencing terms.

$$\hat{\mathbf{Y}}_t = \boldsymbol{\mu} + \boldsymbol{\varphi}_i \mathbf{Y}_{t-1} \quad \text{(Eq. 3.4.1)}$$

As seen from equation 3.4.1, the terms can be defined as follows-

μ – constant

φ_i – parameters used in the estimation

$\hat{Y}_t$ – predicted closing value of indices

Y_{t-1} - lagged values of indices close

ARMA or its variant-based stock prediction models proposed in the prior literature were reviewed in current research. Earlier works (Babu and Reddy, 2014; Hiransha *et al.* 2018) experimented with this model but not on stock indices like banking.

3.4.2.2.Radial basis function (RBF) model

Radial basis function (RBF) is an activation function that computes euclidean distance for approximation function of the target variable, index closing in this case. RBF-based prediction model has been done in stock forecasts (Komo *et al.* 1994, Rout *et al.* 2012, Sheelapriya and Murugesan, 2017). Such a model can also be improved with different learning algorithms outside the current study's scope. The activation function for RBF models can be defined as in equation 3.4.3:

$$\boldsymbol{y(x) = \sum_{i=1}^{N} w_i \; \varphi \left(\left|\left|x - x_i\right|\right|\right)} \qquad \text{(Eq. 3.4.3)}$$

Here y(x) is the output variable, w_i is weighted, and x input data points are input variables from fundamental or technical analysis. The prior literature has indicated that Radial basis functions based neural network gives mixed results in prediction tasks.

3.4.2.3.Vector Auto-regression (VAR)

Due to inherent limitations as observed in ARMA models, the study attempt a VAR system. Vector autoregression-based stock prediction models have been tested in prior works. These include studies by Ibrahim *et al.* 2021; Khansa and Liginlal, 2011 used this model, but tasks related to bitcoin price prediction and security, respectively. Vector autoregression is a powerful method for the stochastic process modeling approach. Here the model captures the linear interdependence among multiple time-series evolving from K variables over any sample period (t = 1, 2 . . . T).

It vectorizes the input data-set. Such variables are collected in a k x 1 vector y_t, as an i^{th} element, $y_{i,t}$, the observation at time "t" of the i^{th} variable.

A *p-th* order VAR, denoted VAR (p), is

$$\mathbf{y_t = c + A_1\, y_{t-1} + A_2\, y_{t-2} + \ldots\ldots\, A_p\, y_{t-p}} + e_t \qquad \text{(Eq. 3.4.2)}$$

In equation 3.4.2, observation y_{t-i} (i periods back) is defined as the i^{th} lag of y, c is k-vector of constants (intercepts). The term A_i is time-invariant ($k \times k$)-matrix, and e_t is k-vector of error terms. We assume that changes in the parameters chosen for current research also impact the economic, financial factors. In essence, the VAR separates the input parameters within the model by vectorizing before estimating the output variable.

3.4.2.4. NARX model

A NARX (Nonlinear Autoregressive exogenous model) is tested due to the high stochastic nature of variables in the data set. The stock's close price is essential as deep learning compared to PCA gave positive results (Gao *et al*. 2016). Stock prediction models that are built using the NARX system in prior literature were reviewed. Some of the prior works were Alkhoshi and Belkasim, 2018 and Gandhmal and Kumar, 2021. While the former study obtained a good convergence rate in the model, the second work tested NARX with feature selection. Also, heterogeneous industries, i.e., multiple sectors, were used, so a limited feature set is got for the financial/banking industry.

$$y\,(t) = f(x\,(t-1),\ldots x(t-d),\, y(t-1),..y\,(t-d)) + \varepsilon t \qquad \text{(Eq. 3.4.4)}$$

The problem defined is to predict series y(t) given d past values of y(t) and another series x(t). Here input attributes are assigned to x(t) columns with y(t) target variable, index closing. The noise/error term is denoted as ε that varies with t as in equation 3.4.4. The thesis now proceeds to implement stated models and analysis of empirical results in the next chapter.

A N A L Y S I S
A N D
D I S C U S S I O N

Chapter 4.

ANALYSIS AND DISCUSSION

In this section, a detailed account of analysis and interpretation of results are provided in this chapter. Simon (1990) argued that in problems of the inherent complexity of modeling, it's vital to decide computational power, temporal information that affects prediction or prescriptions, helpful for the policy level decision making.

4.1. CHARACTERISTICS OF DATA

BSE Bankex: BSE Bankex index is one of the major sectoral indices for banking in India. BSE Bankex currently consists of stocks of nine public and private banks. They are Axis Bank Ltd, City Union Bank Ltd., Federal Bank Ltd, HDFC Bank Ltd, ICICI Bank Ltd., IndusInd Bank Ltd., Kotak Mahindra Bank Ltd., RBL Bank Ltd., and State Bank of India. The past performance of Bankex is shown in Figure 4.1.1. As per the latest data on May 2020, the closing prices of the Federal bank are the lowest in the index with 43.85 rupees and the highest being Kotak Mahindra Bank Ltd. at 1245.4 rupees.

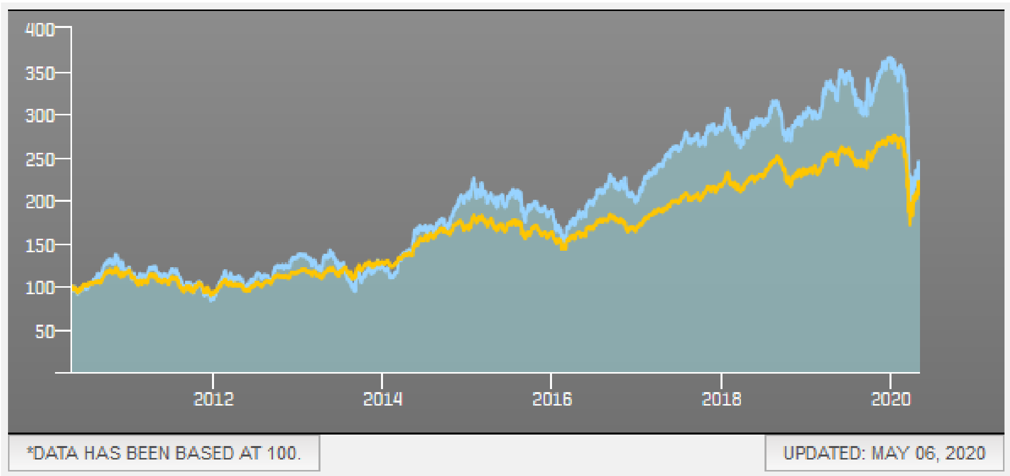


FIGURE 4.1.1: BSE BANKEX DATA SNAPSHOT (SOURCE: ASIAINDEX)

TABLE 4.1.1: COMPARISON WITH BSE SENSEX (SOURCE: ASIAINDEX)

Index name	Index level	Ten-year annual returns
S&P BSE BANKEX (TR) Launch Date: Aug 01, 2006	26,066.04	8.66 %▲
S&P BSE SENSEX (TR) Launch Date: Aug 19, 1996	46,387.57	7.97 %▲

Here, all information is back-tested for an index prior to its launch date, based on methodology that was in place on the start date. Back-tested performance, which is hypothetical rather than real performance, is subject to inherent limitations because it demonstrates hindsight implementation of an Index methodology and collection of index constituents. No theoretical method can take into account all the conditions in the markets in general and the effect of the decisions that may have been made during an index's actual activity. Actual returns will vary from back-tested returns, and be lower than those. (Source: https://www.asiaindex.co.in/indices/equity/sp-bse-bankex)

The BSE Bankex contributes 5.62% of Total turnover in the third position below S&P BSE Fast-moving consumer goods and S&P BSE Finance in all sectoral indices. As of May 2020, the maximum market cap of BSE Bankex is 5,14,252.73 crore INR. Moreover, it has better returns over ten year period than the benchmark index (Table 4.1.1).

NSE NIFTY Bank: The index is designed to reflect the behavior and performance of large and liquid banks. The index comprises a maximum of 12 stocks. The base date of the index is January 1, 2000, with a base value of 1000 points launched on September 15, 2003 (Figure 4.1.2). NIFTY Bank Index is computed using the free-float market capitalization method. NIFTY Bank Index can be used for various purposes such as benchmarking fund portfolios, using index funds, ETF’s (Electronic Traded Funds) products. Table 4.1.2 shows statistics of the NSE Nifty Bank index.

FIGURE 4.1.2: NSE NIFTY INDEX DATA

TABLE 4.1.2: NSE NIFTY BANK INDEX STATISTICS (SOURCE: NSE)

Statistics ##	**1 year**	**5 years**	**Since inception**
Std. Deviation *	38.24	23.03	29.98
Beta (NIFTY 50)	1.20	1.18	1.07
Correlation	0.94	0.91	0.83

*## Based on Price Return Index. * Average daily standard deviation annualized*

TABLE 4.1.3: COMPARISON OF NIFTY BANK WITH GLOBAL BANKING INDICES

Index	**1 Month**	**6 Month**	**YTD (Year-to-date)**	**5 Years***
NIFTY Bank index	1.3	8.1	39.4	15.8
S&P 500 Banks Industry Group Index	2.7	13.4	16.7	N/A
STOXX Europe 600 Bank Index	-2.1	1.0	7.1	2.7
FTSE Bank Index (London)	-0.4	-0.5	7.0	0.5
DAX Bank Index (Dubai)	9.9	16.5	19.0	-4.9
Nikkei Bank Index (Tokyo)	-4.6	6.4	0.4	9.2
Nasdaq Hong Kong Bank Index	0.2	0.5	21.4	N/A

Source: Thomson Reuters, Bloomberg; Note: -Annualized percentage returns, YTD (Year-to-date) as of November-end 20*

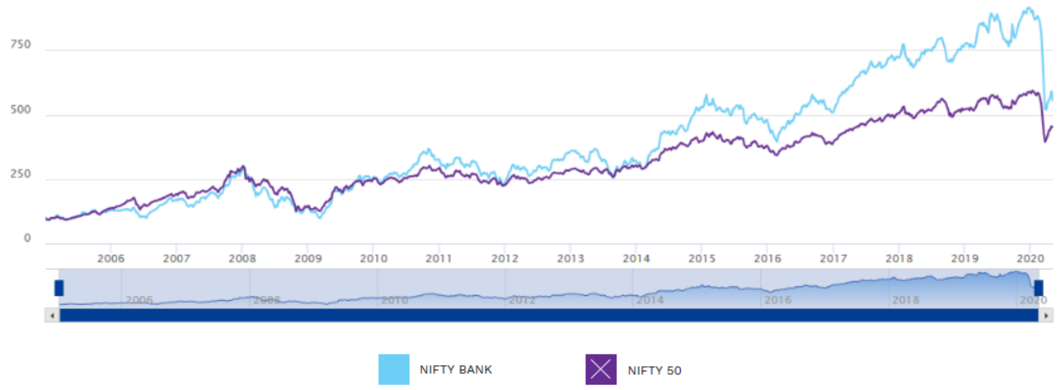


FIGURE 4.1.3: COMPARISON OF NIFTY BANK WITH NIFTY 50 PERFORMANCE

Nifty 50 is the usual comparative benchmark index to track this index (Figure 4.1.3). NIFTY Bank Index outperformed all the major global bank indices on a YTD (Year-to-date) basis, gaining 39.4%. Nasdaq Hong Kong Bank Index also performed better than the other indices, registering 21.4% returns during the same period (Table 4.1.3). The index is calculated using the free-float market capitalization methodology. Also, the effects of risk have been lower compared to NIFTY 50. At the time of rebalancing of shares/ change in index constituents/ change in investable weight factors (IWFs), the weightage of index constituents (where applicable) is capped at appropriate levels (Equation 3.3.1). The weightage of such stocks may increase between rebalancing periods.

*Index Market Capitalization = Total shares outstanding * Price * IWF * Capping Factor (if applicable)* (Eq. 3.3.1)

Here IWF (Investible Weight Factors) is a unit of floating stock expressed in terms of a number available for trading and not held by the entities having strategic interest in a company. Higher IWF suggests a more significant number of shares held by the investors as reported under the public category within a shareholding pattern reported by each company.

4.2. EXPERIMENTAL SETTINGS

Preprocessed data for BSE and NSE is shown in the following figures (Figure 4.2.1 and Figure 4.2.2 having study variables). The original data and smoothed data values of index closing are calculated and visualized using the GRETL tool.

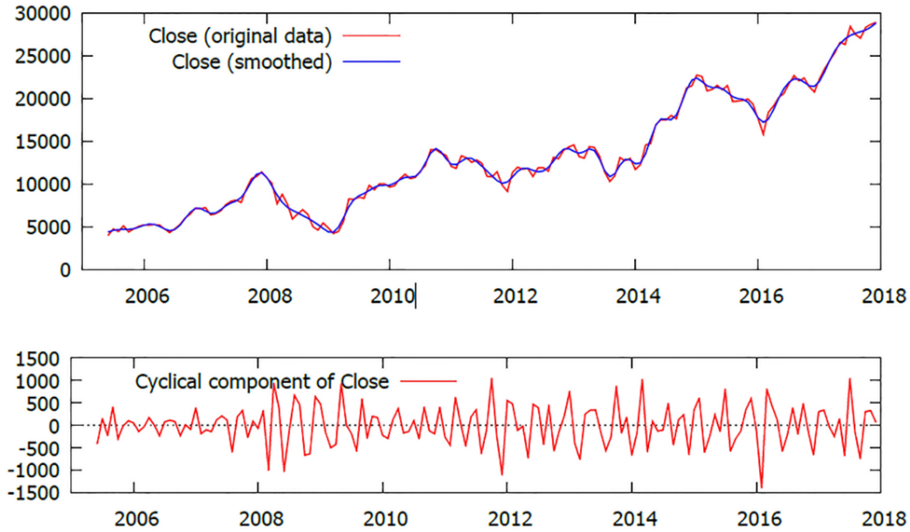


FIGURE 4.2.1: BSE BANKEX CLOSING & CYCLICAL COMPONENT (SOURCE: GRETL)

The cyclical component of close values in Fig. 4.2.1 is red. It indicates 2008 had more volatility due to global crisis events followed by similar variations briefly during early 2012 and post-2016. Figure 4.2.3 shows NSE data with study variables plotted.

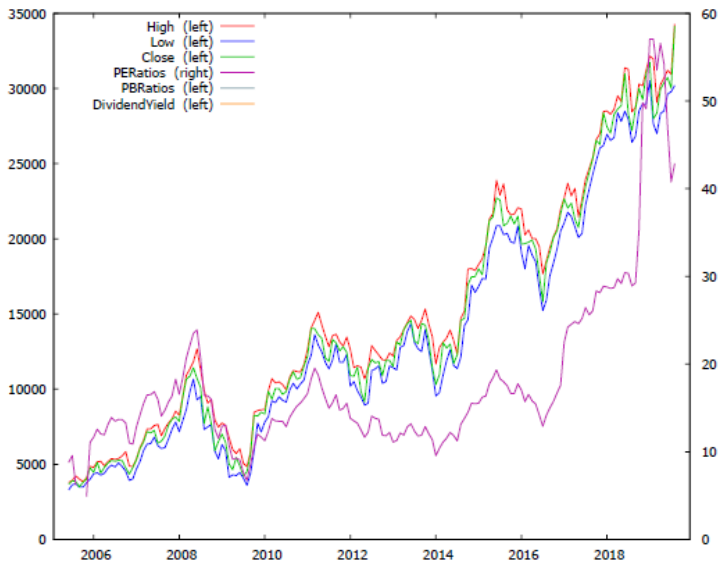


FIGURE 4.2.2: BSE BANKEX CLOSING DATA WITH INDICATORS (2005-2019)

4.2.1. Statistical methods

Before progressing to the model-building or testing process, the variables or parameters required as input to the predictive model need to be assessed. For this, the secondary data from BSE Bankex and NIFTY Bank are used. The data collected is subject to pre-processing methods. These are missing value analysis (MVA), outlier detection, normality check (assuming data follows normal distribution). After this, individual factors are evaluated, i.e., Price-earnings ratio (P/E), Price-book ratio (P/B), Dividend yield, Technical indicators such as Open, High, Low along with constructed indicators like Typical price, Weighted close, Volume Rate of Change and Exponential moving average (EMA).

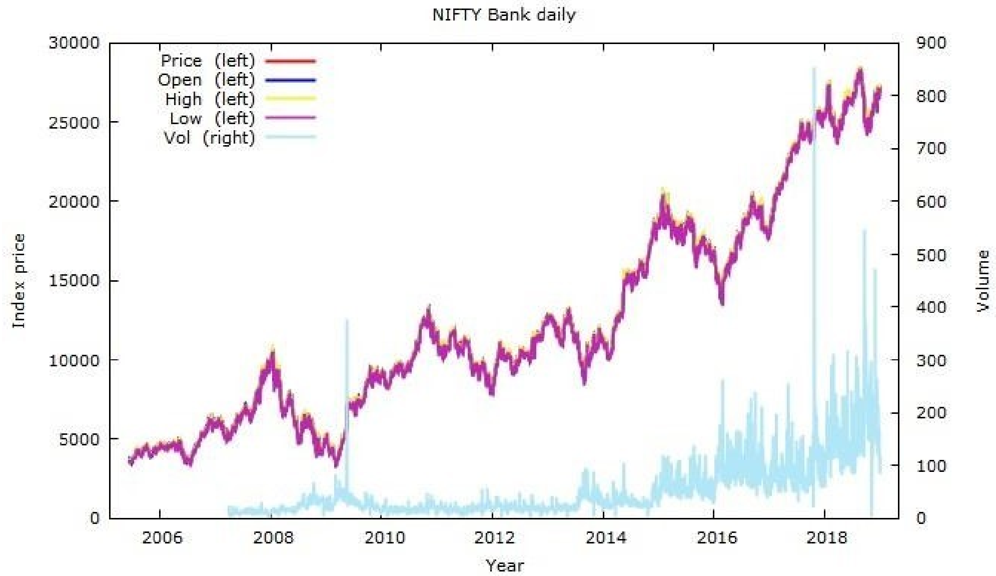


FIGURE 4.2.3: NSE NIFTY BANK DATA WITH INDICATORS (2006-2019)

4.2.2. Pre-processing

The normality of datasets is done using Shapiro-Wilks tests. The results of the test confirmed that observations in both indices data sets were not normal. (W = 0.94639, with p-value = 0.506183) for BSE and (W = 0.929063, with p-value = 0.370283) for NSE. Hence any statistical prediction techniques with reasonable accuracy are of limited scope.

Because of the non-normality of variables involved, modeling methods must be used (Sujatha and Sundaram, 2010b). So, Automatic Linear Modeling (ALM) is used that utilizes regression analysis after the outlier detection, model ensembles, and optimality of variable selections. The first research objective pertains to assessing the predictive importance of variables from fundamental/technical categories, as seen in the earlier chapter. So this objective is being addressed first in the study as follows:

Objective 1: To measure the importance of fundamental and technical indicators that affect the stock index of banking firms in the Indian services sector.

For the objective, a hypothesis is formulated, which is stated below:

- H1 (Null hypothesis): There is no significant difference in the estimation of stock index close of banking sectors using fundamental and technical indicators.
- H1a (Alternate hypothesis): There is a significant difference in the estimation of stock index close of banking sectors using fundamental and technical indicators.

In datasets of BSE, it is found that a negative correlation exists for the typical price to dividend yield (-.7878), with the highest correlation to the P/E ratio (.7887). Hence this suggests negative relation of market closing with dividend yield. Correlation coefficients use the observations 2005:06 - 2019:03. Under 5% critical value (two-tailed) = 0.1524 for n = 166. The AIC (Akaike information criterion) is a metric that assesses the statistical quality of an empirical model and is defined in Equation 4.2.1. AIC is essential to check the efficiency of the model in terms of information available. It is described as:

$$\mathbf{AIC} = \mathbf{2K} - \mathbf{2\,log}(\boldsymbol{L}\hat{}) \qquad \text{(Eq. 4.2.1)}$$

Here, K is the number of parameters in the model. L^ is defined as the maximum likelihood function. As and when the variables 'K' increase, the statistical quality of the model may reduce, and hence proper metric becomes a requirement.

The AIC (Akaike information criterion) with a stepwise forward method produced lower in NSE index data. Still, accuracy reduces over 8% compared to the BSE index with 91.5% (Table 4.3.7). The BIC (Bayesian information criterion) is a similar metric that penalizes errors using log (*n*) instead of *K* from equation 4.2.1. As a formal analysis, the Ordinary Least Squares (OLS) model had technical indicators with N = 151. Here also, the P/B ratio and P/E ratio were most significant with an adjusted R^2= 0.785. To reduce the dimensionality of variables involved in the model, Variable importance analysis (VIA) is computed. The results of VIA are provided in Table 4.2.1. In Figure 4.2.4, the results generated in SPSS are provided.

TABLE 4.2.1: VIA PROCEDURE RESULTS

Target	**Close**
Automatic Data Preparation	On
Model Selection Method	Forward elimination
Information Criterion	214.580

Authors' calculation

Note: The information criterion is used to compare models. Models with smaller information criterion values fit better.

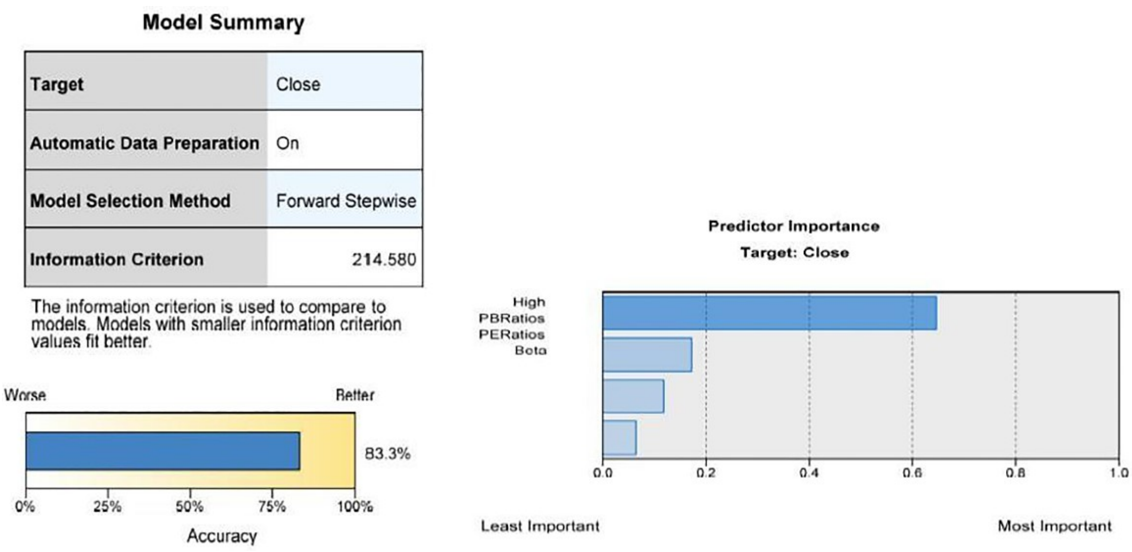


FIGURE 4.2.4: VIA OUTPUT (SOURCE: IBM SPSS)

Forward stepwise selection of model adds the variables to the model, which results in the largest R^2 increase. The VIA procedure shows that more than an 8% accuracy difference exists in predicting BSE Bankex and NIFTY bank indices using ALM (Table 4.3.7). Also, the Beta indicator has the least predictive relevance, so an alternate hypothesis, i.e., H1a, is accepted. So results confirm the difference in the predictive estimation of stock index close for banking sectors using fundamental & technical indicators within BSE and NSE markets.

4.3. MODEL PERFORMANCE RESULTS

As being resolved in objective 1, due to the significant difference between market index estimation within BSE and NSE sectors, the importance of predictive models suited to them becomes an increased requirement for analysts. The second objective is assessing & comparing the prediction performance of statistical models for stock markets.

Objective 2: To evaluate statistical techniques in predicting the stock index price of BSE Bankex and NSE Nifty Bank indices.

i. ARMA model

To achieve the above objective, the ARMAX model is run with observations from yearly data from 2005-2017 (T=13) and is given in Figure 4.3.1.

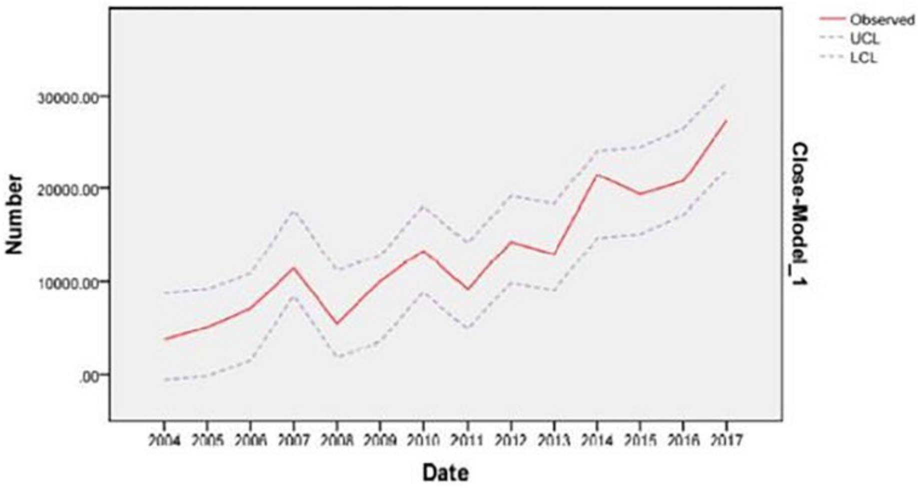


FIGURE 4.3.1: ARMA MODEL OUTPUT (SOURCE: IBM SPSS)

As seen in Table 4.3.1 (ARIMA model results) and Table 4.3.2 (ARMAX), there is significant predictive importance for technical (Open with p=0.0000, High with p=0.0001) but negative co-efficient for the market open (-0.788). The banking indices had low performance in 2008 due to the recession and rebounded in early 2010. Similarly, the dividend yield has significance (p = 0.141) with a negative coefficient (-1595.09). The P/E ratio emerges as the best predictor over the long term (S.E = 151.07). No outliers were detected in the data.

TABLE 4.3.1: ARMA MODEL STATISTICS

Model	**Number of predictors**	**Model Fit Statistics**	**Ljung-Box Q (18)**			**Number of Outliers**
		Stationary R-squared	**Statistics**	**DF**	**Sig,**	
Close-Model_1	5	.976	.	0	.	0

Authors' calculation

ARMA with exogenous inputs (ARMAX) allows multiple predictors to estimate a dependent variable, as seen above. The model performance is reduced at various points, for example, 2007, 2012-13 periods (Figure 4.3.2). It is exciting research to investigate time-specific effects during those time windows on banking sectors. Overall, using five predictors has not affected negatively, and R^2 model fit is 0.976.

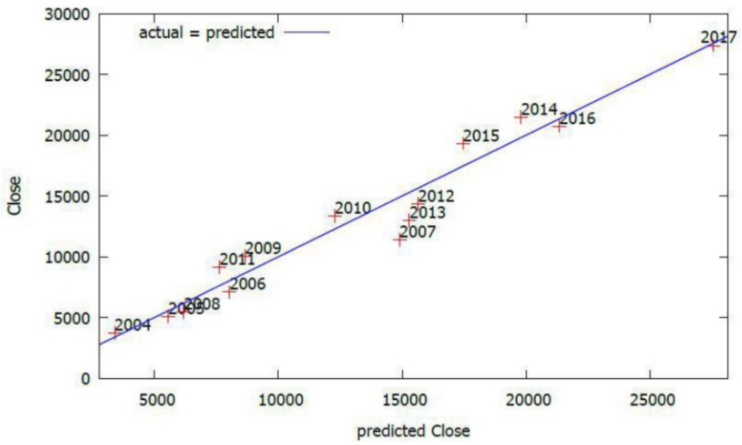


FIGURE 4.3.2: ARMAX MODEL PERFORMANCE (SOURCE: GRETL)

ARMAX model in Table 4.3.2 results show that dividend yield exerts negative relation on market closing. At the same time, better opening performance affects indices closing as well. The standard error is highest in estimates using Beta as a predictive factor.

TABLE 4.3.2: ARMAX MODEL RESULTS

Particulars	**Co-efficient**	**Std. Error**	*z*	**p-value**
const	650.767	440.246	1.4782	0.394
P/E Ratios	170.314	151.070	1.1274	0.2596
P/B Ratios	36.1980	1041.76	0.0347	0.9723
DividendYield	-1595.09	1084.74	-1.4705	0.1414
Beta	939.052	1554.49	0.6041	0.5458
Open	-0.788923	0.151427	-5.2099	0.0000
High	0.68108	0.171579	3.9691	0.0001
Low	0.481001	0.196719	2.4451	0.0145
Mean dependent var	1819.491		S.D. dependent var	4261.305
Mean of innovations	0.000000		S.D of innovations	620.2946
Log-likelihood	-102.0387		Akaike criterion	220.0775
Schwarz criterion	224.5971		Hannan-Quinn	219.1485

Note: Dependent variable (1 – L) Close

Authors' calculations

ii. VAR system (Model summary source: GRETL)

VAR system, lag order 12

OLS estimates, observations 2006:06-2017: 12 (T=139)
Log-likelihood = 1150.3613
Determinant of covariance matrix = 903548.1
AIC = 16.7822
BIC = 17.1200
HQC = 16.9194
Portmamteau test: LB (34) = 14.2679, df = 22 [0.8917]
Equation 1: Close
Heteroskedasticity-robust standard errors, variant HC1

As can be inferred from Table 4.3.4 and the model performance chart (Figure 4.3.3), only the P/E ratio has a positive coefficient in the prediction estimate (66.28) and the least standard error (30.15).

TABLE 4.3.3: VAR SYSTEM MODEL

Mean dependent var	1372.50	**S.D. dependent var**	6280.591
Sum squared residue	1.26e+08	**S.E of regression**	1010.486
R-squared	0.976928	**Adjusted R-squared**	0.974114
F(15, 123)	455.5764	**P-value (F)**	4.9e-100
Rho	0.12648	**Durbin-Watson**	1.971584

Authors' calculation

The model shows that Close with 1 lag has a 99% confidence level and positive co-efficient (1.07). This factor, more significant than the P/E ratio, indicates that the banking sector market exhibits memory. Other lagged values of any variables did not show statistical significance. Even though this is the case, the standard error remains high at 1010.48 (Table 4.3.3)

TABLE 4.3.4: VAR MODEL RESULTS

	coefficient	**std. error**	**t-ratio**	**p-value**
Const	1760.43	2186.70	0.8051	0.4223
Close_1	1.04740	0.102642	10.20	4.36e-018 ***
Close_2	-0.191345	0.130916	-1.462	0.1464
P/E ratios	66.2848	30.1599	2.198	0.0298 **
P/B ratios	-591.929	392.147	-1.509	0.1337
Dividend Yield	-517.850	866.328	-0.5978	0.5511

Authors' calculation

**** Statistically significant at 99% confidence level*

*** Statistical significance at 95% confidence level*

VAR model creation using the GRETL software as default procedure also runs OLS (Ordinary least squares) analysis. Results are provided in summary above. The BIC is higher than AIC, indicating that more penalization of errors has occurred among input predictors.

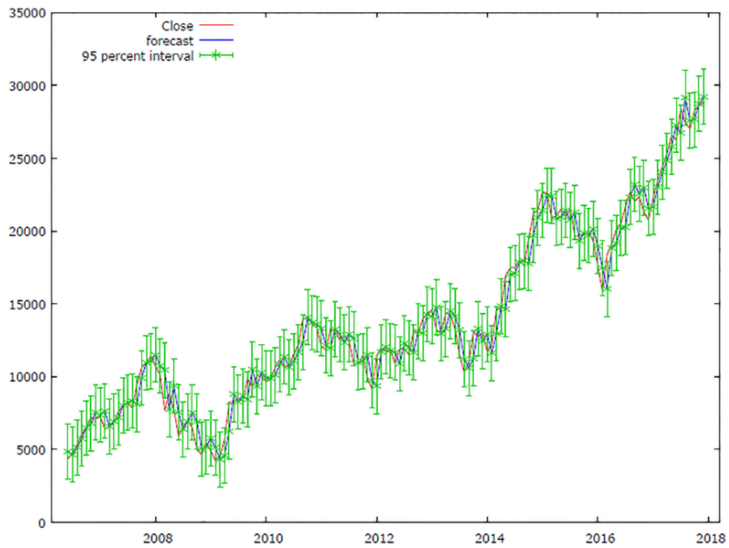


FIGURE 4.3.3: VAR MODEL PERFORMANCE (SOURCE: GRETL)

Results indicate that the model can utilize high-frequency tick data (15-minutes window) as adopted by Selvamuthu *et al*. 2019 using the ANN model.

Due to the inherent properties of statistical models from results, the efficacy of using an artificial neural network model for market estimation is being addressed next. Hence the findings show that the VAR model is suited in the prediction of NSE, ARMAX for BSE prediction. Experimental results indicate the scope to use high-frequency data as well. Obtaining high-frequency data is challenging for two reasons. High-frequency trading (HFT) is still at a nascent stage in India. Second, algorithmic trading lacks regulatory standards. Third, this data is limited by specialized financial information providers like Bloomberg, requiring subscription or fees. Artificial neural networks have successfully been shown in prior studies to overcome non-linearity etc. The proposed framework requires augmenting statistical model results to process stochastic stock market information using neural networks.

Objective 3: To compare the different artificial neural networks in stock index prediction based on forecast accuracy. Various models are developed and tested to explore this objective and gather insights on various artificial neural network architectures.

iii. Multilayer Perceptron (MLP) model

Among all the ANN models from the literature, multilayer perceptron models have been shown to have high performance for both prediction and classification tasks. Hence in the current study, the MLP model is tested using the hyperbolic tangent activation function at hidden nodes giving lower errors. This activation function is defined in equation 4.2.3:

$$\tan h(s) = \frac{\sin h(s)}{\cos h(s)} = \frac{e^{\alpha s} - e^{-\alpha s}}{e^{\alpha s} + e^{-\alpha s}} \quad \text{(Eq. 4.2.3)}$$

The results of implementing the MLP model are explained in Tables 4.3.5 and 4.3.6.

TABLE 4.3.5: MLP MODEL NETWORK INFORMATION

Input layer	Factors	1	P/E Ratios
		2	P/B Ratios
		3	Dividend yield
		4	Beta
		5	High
Hidden layers	Number of Units[a]		51
	Number of Hidden Layers		1
	Number of Units in Hidden layer 1[a]		4
	Activation Function		Hyperbolic tangent
Output Layer	Dependent variables = 1		Close
	Number of units		1
	Rescaling method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

Authors' calculation

a. *Excluding the bias unit*

TABLE 4.3.6: MLP MODEL SUMMARY

Training	Sum of Squares Error	.002
	Relative Error	.000
	Stopping Rule Used	Training error ratio criterion (0.001) achieved
	Training Time	0:00:00.00

Dependent variable: Close

Authors' calculation

Table 4.3.5 indicate the architecture of the MLP model. The Multi-Layer perceptron (MLP) with backpropagation algorithm outperforms other models for NSE index forecasting. Results showed the sum of squares error (SSE) = 0.002 and relative error value 0.000 proves strong applicability. The model also identifies the P/B ratio as the most important with the least important Beta (Table 4.3.7). Findings comply with the finding of Wu and Hu, 2012. The experiment used a hyperbolic tangent activation function with {51-4-1} architecture. The training time is 0.00, which is minimal to setup the predictive model (Table 4.3.6)

iv. Radial basis function (RBF) model

The predictive importance estimation calculated by Radial basis functions indicated its poor performance efficiency. The sum of squares in BSE and NSE predictions are SSE=1.64 and SSE= 3.66, respectively (Table 4.3.7). Findings are a contradicting result comparing RBF architecture against a finding by Komo *et al*. 1994 that produced a superior prediction of stock forecasts of the Dow Jones index.

The earlier research findings indicate a Bayesian algorithm to improve the predictive accuracy (Sheelapriya and Murugesan, 2017). Still, its validation using data from the Indian markets is an open research problem. Hence empirical results indicate MLP models are far superior to RBF among the artificial neural network architectures.

This inference is the same with the case of studies that experimented with data from S&P and Dow Jones industrial average (Rout *et al*. 2012). Hence findings indicate that radial basis functions in predictive models generate large errors, accuracy and lower model convergence for forecasts.

v. NARX model

In MATLAB, *ntstool*, a neural network tool wizard, is run in the simulation experiment procedure.

A total of 1328 time-steps of datasets is used. Training used 70% of data while testing and validation are 15% each. The algorithm of Levenberg-Marquardt was implemented, with no. of hidden neurons ten and delay units *d*, set to 1.

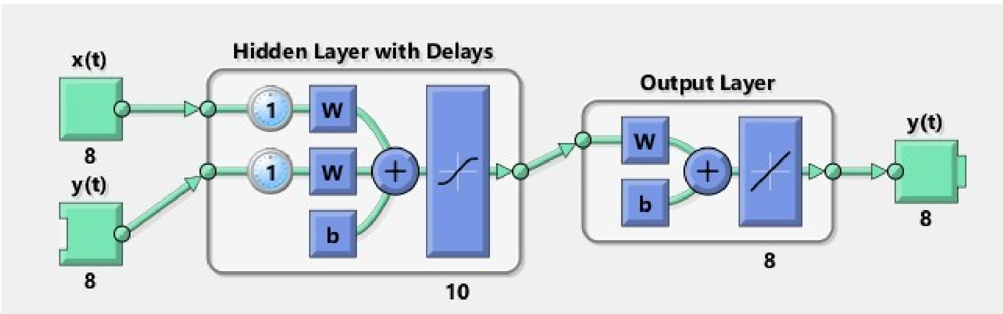


FIGURE 4.3.4: NARX MODEL ARCHITECTURE (SOURCE: MATLAB)

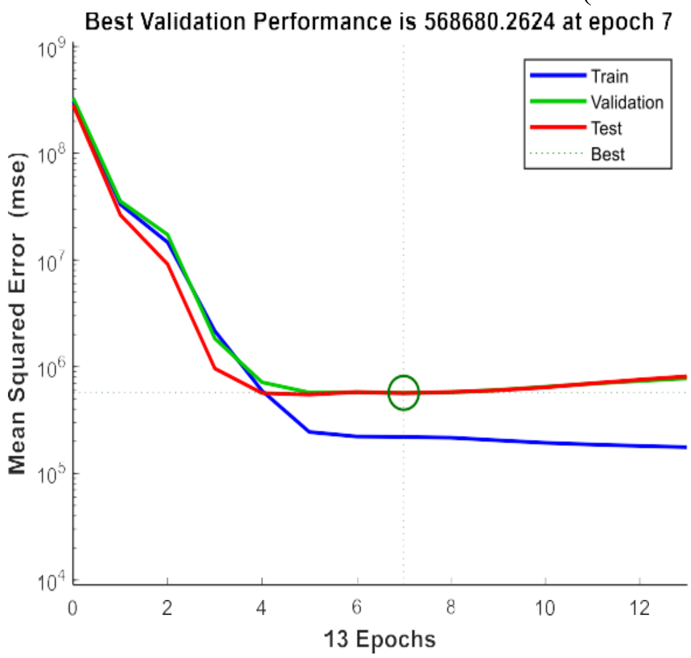


FIGURE 4.3.5: NARX PERFORMANCE VALIDATION (SOURCE: MATLAB)

The computed technical indicators are calculated using the same method in prior literature (Gao *et al*. 2016). The NARX system is shown in Figure 4.3.4.

Negative correlation is observed for typical price to div. yield (-.7878) with highest correlation to P/E ratio (.7887). Correlation coefficients uses the monthly observations 2005:06 - 2019:03. Under 5% critical value (two-tailed) = 0.1524 for n = 166. From NIFTY Bank dataset reports correlation analysis. The coefficients, use observations in 2007-03-29-2019-06-10 (missing values skipped). With 5% critical value (two-tailed) = 0.0357, n = 3014**.**

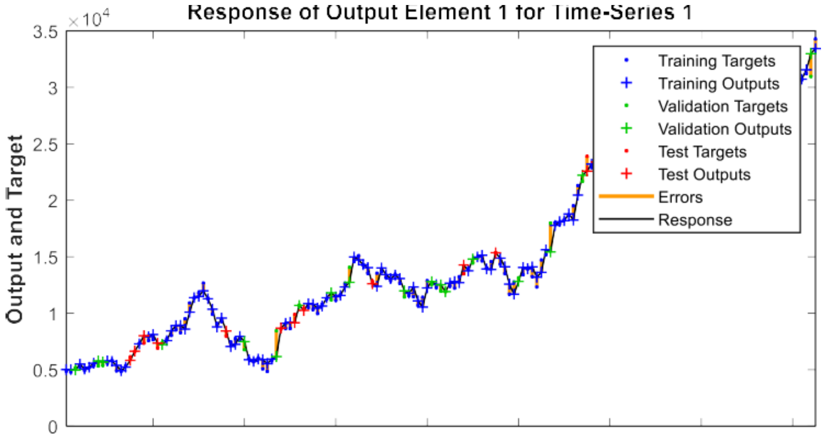


FIGURE 4.3.6 : TIME SERIES RESPONSE (SOURCE: MATLAB)

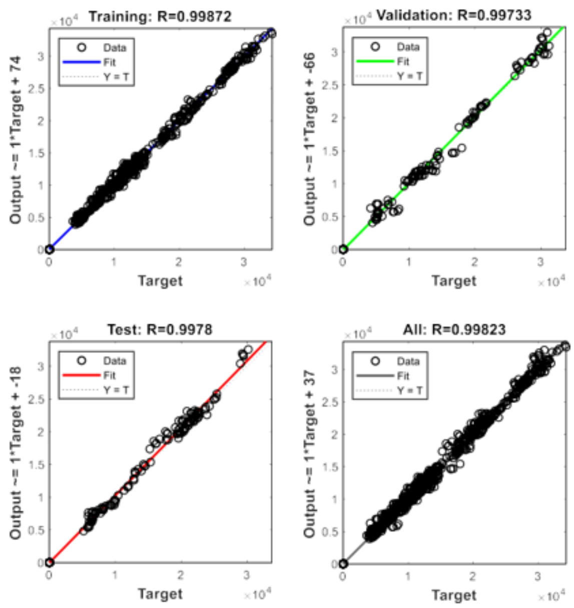


FIGURE 4.3.7: REGRESSION OUTPUTS (SOURCE: MATLAB)

In the experiment, the total number of data points was 1494, including all variables. Technical indicators used in the study followed are listed below:

1. Typical Price (TP): Uses High (Hi), Low (Lo) and close price (Cl) given as below:

$$\mathbf{TP} = \frac{\mathbf{Hi+Lo+Cl}}{\mathbf{3}} \qquad \text{(Eq. 4.2.6)}$$

2. Volume Rate of Change (VROC): Uses 14 day period is taken and volume variable

$$\mathbf{VROC} = \frac{\boldsymbol{Vo-Vo_{t-14}}}{\boldsymbol{Vo_{t-14}}} \qquad \text{(Eq. 4.2.7)}$$

3. Exponential Moving Average (EMA): This is calculated as

$$\boldsymbol{MA = [Price - Previous\ EMA * \left(\frac{2}{n+1}\right) + previous\ EMA]} \quad \text{(Eq. 4.2.8)}$$

Here n = 20 days is assumed for the input data sets as seen in prior studies.

4. Weighted Close (WC): Calculated by the equation as below:

$$\mathbf{WC} = \frac{\mathbf{Hi+Lo+Cl*2}}{\mathbf{4}} \qquad \text{(Eq. 4.2.9)}$$

The model could generate accurate prediction estimates using 37 monthly data points or 3.5 yearly data sets. The higher error rate observed in 40-110 time-steps can be reduced with daily datasets (Figure 4.3.6). The model iterations stopped at 13 epochs. In epoch seven, the best performance instance after retraining and training regressions shows validation tends to reduce R^2 (Figure 4.3.7). However, the simulation results offer higher R^2 performance while requiring lesser model iterations and time, leading to convergence higher than Alkhoshi and Belkasim (2018). Comparative performance analysis is done to gain a broader understanding of strengths and shortcomings using various experimented approaches. Predictive performance is measured using R^2 model fitness, Sum of squares (SSE), and relative error observed and provided in Table 4.3.7. One of the challenges in performance evaluation is the lack of standard benchmark datasets, which are possible in other research areas like computer vision. Due to the absence of empirical benchmarks, statistical models are hard to formulate. From the findings, the MLP with backpropagation is a better predictive model for NSE prediction. At the same time, NARX also exhibits similar accuracy and lower error rates for BSE. Hence, from the analysis, the alternate hypothesis H2a is being accepted. Hence, ANN models have a substantial advantage for predictive relevance in stock index prediction.

Objective 4: To create an optimum artificial neural networks-based prediction model for forecasting the stock price index in this sector.

Since the objective is exploratory in empirical nature, the study adopted multiple approaches. By such strategy, different aspects of predictive modeling techniques, for ex: qualitative factors, information system adaptations and management decision support system success-related literature, get amalgamated (Kolora and Pandey, 2020; Qiu and Song, 2016). Subsequently, the results are to be integrated into a broader framework of intelligent decision support system functionalities. The study used daily search query data to measure the individual's interest in the aggregate stock market. It found that investors' attention to the stock market rises in times of high market movements. Additionally, data mining tools such as Google trends and algorithms have been shown to extract predictive information of economic indicators from public internet trends (Choi and Varian, 2012). They show that the ARIMA model using relevant Google Trends variables outperforms models that exclude such predictors by 5% to 20%. For addressing research gap 3, the study looks at whether the stock price search interest has a significant impact through predictive relevance in the Indian stock market. Thus, the present research attempts to integrate Google search query volume trend-based prediction and its impact on the Indian stock market closing. For this, Google Trend data for terms "*stock price*" is retrieved from a global search. Along with this, the secondary data of the Indian market closing and other markets are compared.

Google Trends algorithm:

This qualitative factor is measured using the keyword "*Stock price*." Numerical values represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity of the term. A value of 50 means that the term is half as popular. Likewise, a score of 0 means the term was less than 1% as popular as the peak. The algorithm reports a query index. The query index starts with the query share, i.e., the total query volume for the search term in a given geographic region divided by the total number of queries in that region at a point in time. The query share numbers are then normalized so that they started at 0 on January 1, 2004. Numbers at later dates indicated a % deviation from query share on 01-01-2004.

The conceptual model developed is in Figure 4.3.8. To answer the investigate questions, the hypothesis is:

- H2 (Null hypothesis): There is no statistically significant relation of the global stock price search interest on the Indian stock market index close.
- Alternate hypothesis, H2a: There is a statistically significant relation of the global stock price search interest on the Indian stock market index close.

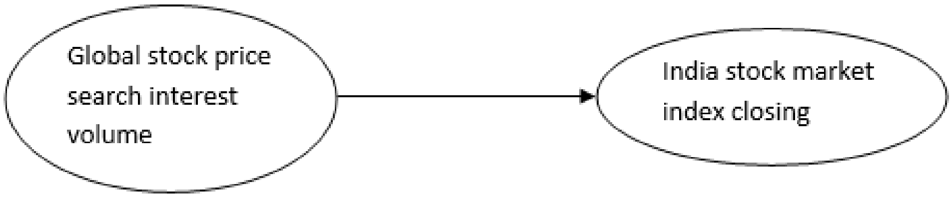


FIGURE 4.3.8: HYPOTHESIZED MODEL (AUTHOR'S OWN)

The exploration and empirical testing of this model is explained in the next section. To achieve qualitative factors affecting the stock index, data collection from worldwide markets, including India, is done. The highest market capitalization-based country-specific indices are chosen as per Bloomberg data. Such has been done since search query returns information of these specific index values in these markets.

In USA- DJIA (Dow Jones Industrial Average), India- BSE 30 (Bombay Stock Exchange), Singapore- STI (Strait Times Index), Hongkong- HSI (Hang Seng Index), Canada- TSX (Toronto Stock Exchange) which are all benchmark indices in respective countries and highest market capitalization. Though data is from 2012-2017 (Figure 4.3.9), future models can acquire such information by real-time streaming also.

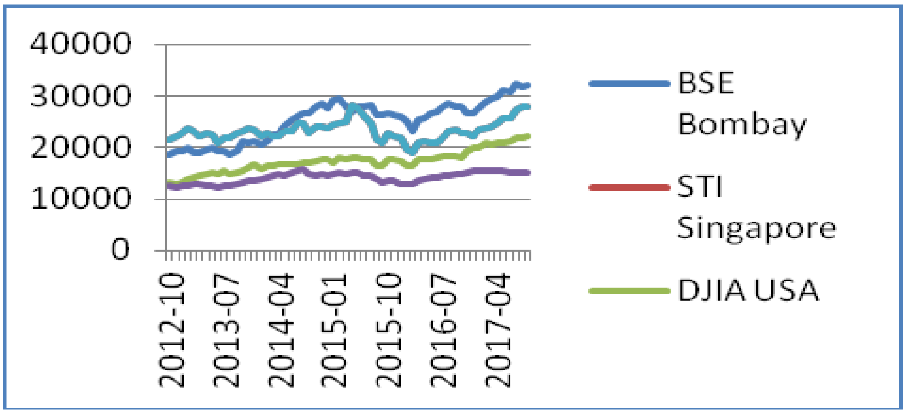


FIGURE 4.3.9: GLOBAL STOCK INDICES PERFORMANCE (SOURCE: FINANCE.GOOGLE.COM)

Internet search patterns: The web search queries and trends data have been collected from Google Trends, including price and query location (Figure 4.3.10). Singapore has the highest search interest while India occupies the 5th position (Table 4.3.8). The result could be due to high fintech proliferation comparing to other nations. As the figure below indicates, a sharp increase in stock search interest occurs in 2015-2016.

FIGURE 4.3.10: STOCK PRICE SEARCH POPULARITY TREND (SOURCE: GOOGLE TRENDS)

TABLE 4.3.8: DEMOGRAPHIC DISTRIBUTION OF COUNTRIES

Country	**Stock price: (1/1/04 - 9/13/17) Search interest query volume**
Singapore	100
United States	75
Canada	68
Hong Kong	53
India	43
Japan	1

Source: Google Trends

To know if the relationship exists between search interest and markets is spurious or causal, correlations analysis is done. Here primary data sets of stock indices prices to search query volume are obtained from the Google trends algorithm. Statistical tests used Spearman Rho, a non-parametric method, and results are given in Table 4.3.9.

TABLE 4.3.10: ANOVA

Model	Sum of squares	df	Mean Square	F	Sig.
1 Regression	558900274.2	1	5598900274.2	81.169	0.000[b]
Residual	399367079.2	58	6885639.297		
Total	958267353.4	59			

Authors' calculation

a. *Dependent variable: BSE Bombay*
b. *Predictors: (Constant), Stock price (Worldwide)*

TABLE 4.3.11: BIVARIATE REGRESSION COEFFICIENTS [a]

Model	Unstandardized Coefficients		Standardized coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error	Beta			Lower bound	Upper bound
1 (Constant)	16675.913	1009.116		16.525	.000	14655.947	18695.879
Stock price (Worldwide)	156.419	17.362	.764	9.009	.000	121.665	191.172

Authors' calculation

a. *Dependent Variable: BSE Bombay*

TABLE 4.3.12: MODEL SUMMARY[b]

Model	R squared	Adjusted R-squared	Std. Error of estimate
1	.583	.576	2624.05017

Summary statistics quantify whether searching affects markets verified by the ANOVA (Analysis of Variance) test (Table 4.3.10). Here, the popularity of the global stock price quest for the BSE 30 closing index showed a Pearson correlation value of .764. The non-parametric Spearman correlation value (.792) between the global price search index and BSE stock closing. Both values are statistically significant in two-tailed hypothesis tests at .01 (99 percent confidence level). Also noteworthy is that India shows a considerable correlation to DJIA and global trends among Asian countries based on search volume (Table 4.3.9). Also, Japan is the country that has the least measured stock price search behavior. The findings support evidence that the online stock price search interest is statistically related to closing the USA, India, and Canada stock market index. The Indian BSE stock market is also closely correlated with the US-based index (.916) of DJIA (Dow Jones Industrial Average).

The bivariate regression analysis of stock search price interest on BSE market closing denotes the $R^2 = 0.583$ for, N=60. The degrees of freedom, df = 58, check whether global search popularity affects the Indian stock market BSE index (Table 4.3.11). Assuming a linear relationship, it shows the determination coefficient of 57.6 percent in the BSE index closing from the global stock price search (Table 4.3.12). The t value 16.525 > 12.706, (t-critical value) for a 2-tailed test at α=.05, thereby alternative hypothesis H2a is accepted. So as per findings, an increase in global stock price search interest does impact Indian market closing.

4.4. DISCUSSION

In this subsection, I elaborate on the findings gathered after simulation and experimental works. Each of the research questions posed with objectives and its solutions from theoretical and empirical perspectives is considered.

4.4.1. Major factors affecting banking index

This subsection interprets the obtained results and findings of objective 1. As noted from multiple simulation experiments and model estimations, there have been mixed results regarding predictive indicators between markets. One of the critical findings implies that Beta, which indicates volatility in the market, has relatively least predictive importance in stock index prediction for the banking sector in NSE. Other financial ratios are variables that have more impact on the closing price. In the case of BSE, the dividend yield has rich predictive information. Such a result is contrary to general literature within core finance and econometric research.

Nevertheless, systematically developed predictive modeling excluding volatility can accurately estimate the bank stock index represented in BSE and NSE markets for day-week forecasts. Such estimations do not take into account any major events that are deviating from usual markets. Also, the significant predictive indicator for BSE is the Price-Earnings ratio, which emphasizes the extent of dependence of the factor in index closing. In the NSE index for banking, the price-book ratio is a crucial determinant factor on the indices closing performance. The stakeholders can consider and utilize these findings in analyzing market trends and forecasts.

4.4.2. Efficacy of statistical models in prediction

In objective 2, the research aimed to build and test few predictive models using statistical approaches. To this end, various models experimented with the secondary data collected. The findings indicated that while ARMA variant models exhibited good stationary R^2 and fitted data. But the lack of predictive indicator importance estimation values and outliers in the dataset came as a slight hindrance. The only marginal difference was observed for predictive accuracy between BSE and NSE.

While this is the case, usage of fundamental and technical variables using statistical models is evident from the standard error residuals. The p-value of indicators was also calculated with the Akaike information criterion to ensure the quality of the model. There was a negative co-efficient among the two factors used in the model, i.e., Open and Dividend yield, technical and fundamental indicators.

The results obtained from VAR models suggest significantly higher accuracy accommodating lagged values as input variables. But even here, the standard error (S.E) of regressions returns higher values that show possible weakness in long-term forecasts of more than 1-2 month window. Besides the earnings ratio as a positive factor, all other input variables negatively influenced the index closing. While a 95% confidence interval in estimation values, there exists more scope for improvement. These are the usage of high-frequency data in the VAR model. It remains a future scope as it requires a dedicated terminal provided in subscription/pay mode by providers like Bloomberg.

4.4.3. Comparative analysis of neural network models

After the simulations, performances of neural network architectures are measured and compared using standard metrics. The values of Root mean squared error (RMSE) were calculated. The findings showed that the Multilayer perceptron model (MLP) has RMSE=0.002 for NSE NIFTY BANK prediction and 0.005 for BSE Bankex prediction. Earlier similar models proposed in the literature for Bankex prediction by Balaji *et al.* 2018 had 71.5% accuracy. Still, they implemented more complicated hardware architecture in the model requiring more setup complexity, processing time. The results of Birau *et al.* 2018 indicated Bankex absorbed the global financial crisis well. The factor of the Beta indicator not having a significant impact on the stock index closing is in line with the result. Mohapatra and Misra, 2019 found out that price-earnings influence Indian banking stocks and price-books with the highest t-ratio and p-value <0.000 quarterly data within the 4-factor model. Future ANN models can be extended for estimating the accuracy of the 4-factor model as done by Charumathi and Suraj, 2014.

4.4.4. An optimum neural network model for stock index prediction

Earlier works like Hu *et al*. 2018 and Qiu and Song, 2016 obtained good predictive results because Google trends data improved prediction model accuracy. Though statistically significant predictive power was identified, there is less benefit due to lower accuracy in such qualitative data to the prediction framework. These findings are in line with the studies like Nardo *et al.* 2016. The lower efficiency of google Trends in prediction model results aligns with other studies in emerging markets like the Korean stock index (Pyo *et al*. 2017). Hence other optimization methods like technological upgrading, deep learning etc. needed in frameworks (Kelotra and Pandey, 2020)

4.4.5. Implementation of intelligent decision support systems in the sector

The banking sector exhibits a strong association with the technological sector and is reflected in market performance. So, innovations and advancements from information technology require an implementation roadmap in the sector. For this extended analysis of the latest technology areas is done. Blockchain is a decentralized ledger that offers high data security and privacy. As a disruptive technology, adoption and implementation in India are also being encouraged by industry and government. Following the Securities and Exchange Board of India (SEBI), an advisory committee, i.e., Committee on Financial and Regulatory Technologies (CFRT), understood potential blockchain application-specific in the stock market. Accenture Inc. (2016) report suggests that in a survey of 32 bank executives' interviews, 9 out of 10 opined that banks focus on technology adoption. In the Economic Times op-ed article, the main benefits of blockchain technology as far as the stock market is concerned are few; first *transparency*, *interoperability* and *trust* within the market system. The second is logistic advantages. For example, reducing to less than three days for transaction settlement between traders, brokers, regulators and exchange itself. Due to this, optimal resource usage is brought in through automation and decentralized operations. A similar aspect was studied by Liu and Yu, 2018, Pop *et al*. 2018. The current study attempted to evaluate the efficacy of blockchain not limiting to stock market operations. Before a conceptual model can be designed, prior works on blockchain research are surveyed.

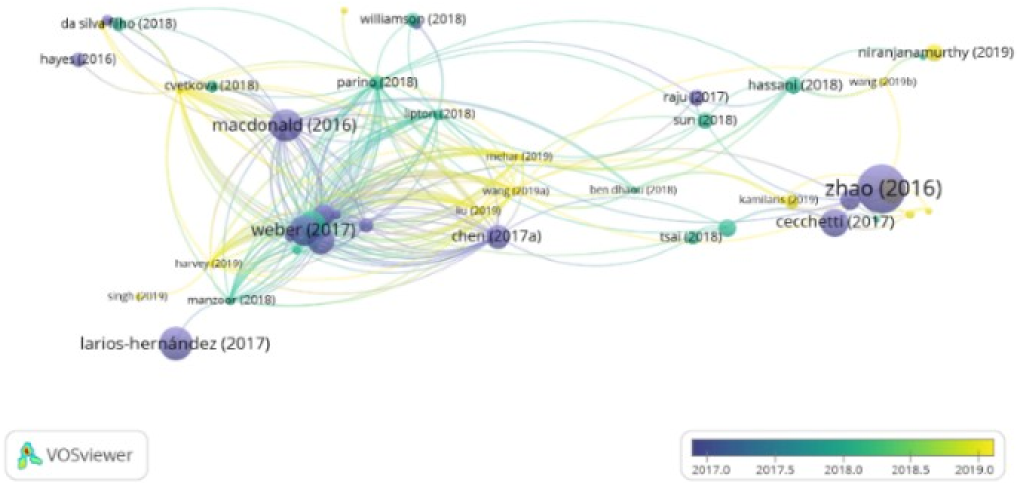


FIGURE 4.4.1: BCAD OUTPUT (SOURCE: VOSVIEWER)

For this, a Bibliographic coupling analysis of documents (BCAD) was done using VOSviewer software (Figure 4.4.1). The published literature from Scopus (largest database) and Web of science was analyzed, including 72 articles. The yellow nodes represent recent research and dark blue nodes older published articles. More related studies exclusively focusing on banking are Guo and Liang (2016) and Xie *et al*. 2020. Guo and Liang (2016) have been the most influential paper on the topic from Scopus, in which the banking industry was comprehensively investigated. The other influential article is Gombe *et al*. 2018 that addressed blockchain and other advancements under the broader framework of FinTech.

Several theoretical underpinnings both from information systems and management literature formed the basis of the majority of works. These include Technology Acceptance Model (TAM), Status quo bias theory and nudge theory, respectively. Also, heterogeneity exists in a type of banking network in each economy. Under tenets of Adaptive structuration theory (AST), i.e., for ex., in developing economies like Brazil, various entities like retail/post office banks etc. require appropriations of multiple actors beyond the IT artifact to assess the technology, position, usage, social, and policy level effects (Leonardi *et al.* 2016).

FIGURE 4.4.2: BLOCKCHAIN TECHNOLOGY IMPACTS (ARJUN AND SUPRABHA, 2020)

In Figure 4.4.2, the far-reaching impact of the adoption of blockchain technology on industries, specifically in the banking context, are outlined. This framework was put together after extensive literature review on blockchain applications within the banking and financial services industry. The main components were delineated from the core research and development of blockchain systems. A business model is fundamental in industry. It is also affected by standards and compliance, industry/market policies, and legal requirements. The detailed list of scientific sources of references that constitute the framework is given in Appendix-I (Tables 4.4.1 and Table 4.4.2). The market-specific factors play a significant force in the effectiveness and success of management information systems. Developments in blockchain affect the overall business model of firms, the standards and compliance, policies of the company and even the legal aspects of their existence. In figure 4.4.3, the hierarchy of requirements that necessitate embracing blockchain into the organizational workflow is identified. The top-down approach can shift to adapt the information system with organizational readiness in terms of company vision, mission, and objective to align with such technological upgrades.

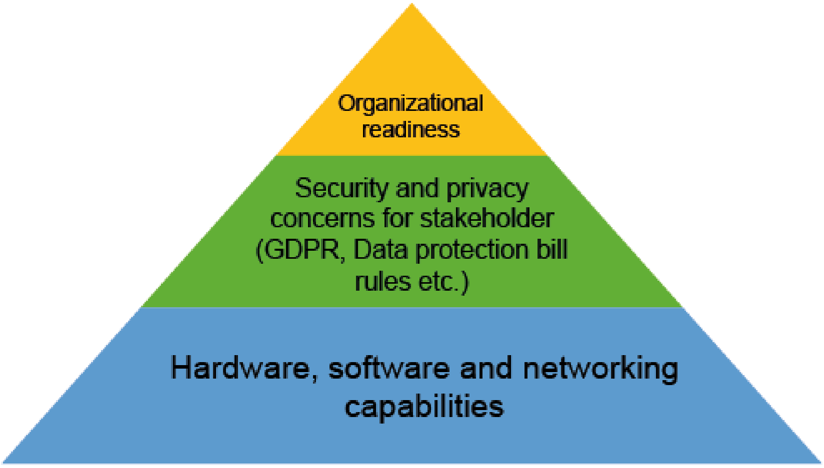


FIGURE 4.4.3: TECHNICAL REQUIREMENTS HIERARCHY (ARJUN AND SUPRABHA, 2020)

Next, as in the current digital age, firms must guarantee the security and privacy of the data about stakeholders (investors, employees, customers, etc.) whenever blockchain or similar technologies are implemented.

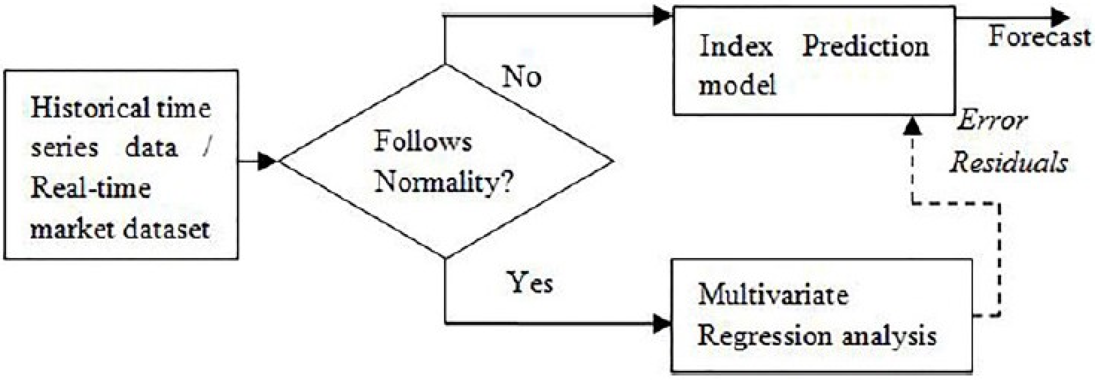


FIGURE 4.4.4: STOCK PREDICTION FRAMEWORK (ARJUN AND SUPRABHA, 2019)

An overall framework to improve stock market prediction systems was designed and shown in Figure 4.4.4. Specific use cases at each stage of the information systems development cycle are identified and reviewed (Table 4.4.1 and Table 4.4.2- Appendix I). Finally, the results of the formulated objectives and hypothesis of the research are outlined in Table 4.4.3.

TABLE 4.4.3: RESEARCH OBJECTIVES AND RESULTS

Research objective	**Results**
1. To measure the importance of fundamental and technical indicators that affect the stock index of banking firms in the Indian services sector.	The price-to-earnings ratio (P/E) has the highest predictive performance, with Opening having a significant effect on market closing. BSE and NSE have major variations in terms of linearity in data. Statistical estimation quality higher in NSE, but accuracy gets lower. **Hypothesis 1 alternate is accepted**.
2. To evaluate statistical techniques in predicting the stock index price of BSE and NSE.	The vector-autoregressive model has the highest prediction accuracy in NSE and ARMAX for BSE modeling. **Hypothesis 2 alternate is accepted.**
3. To compare the different artificial neural networks in stock index prediction based on forecast accuracy.	Multilayer perceptron model (MLP) performs better than Radial basis function in stock prediction tasks. NARX gives comparable accuracy as MLP in the BSE stock index forecast.
4. To create an optimum artificial neural networks-based prediction model for forecasting the stock price index in this sector.	Optimization of neural network model requires the diffusion of information sources. Technologies like blockchain support the paradigm of FinTech. Stock price search interest has significant predictive power in BSE market trends though it’s less robust and error-prone. **Hypothesis 3 alternate is accepted.**

C O N C L U S I O N S

Chapter 5.

CONCLUSIONS

In this chapter, the summary of results, conclusions are drawn from the research and explained. The limitations of study and implications of work in technical, management and organizational dimensions are elaborated. Finally, the future scope and directions of research are provided. Specifically, the contributions in terms of technical, financial and management perspectives are discussed as well.

5.1. SUMMARY OF FINDINGS

Based on the analysis and interpretation, the four objectives had been fulfilled in the research study. The first objective was to quantify major factors affecting stock market closing in bank indices was carried out. The findings indicated that the P/E (Price-earnings) ratio is the most crucial factor that affects BSE and NSE bank sector indices. Implications are that the company’s earnings announcements make a short-term impact. Consistent firm performance is reflected in share price, which affects the investors' decision-making. The findings from the chapter point to the fact that fundamental and technical indicators essentially capture long-term and short-term trends respectively within banking sector indices as generally perceived. While events and news make a significant market short-term impact, long-term evaluation by investors depends on consistent performance in terms of high returns, lower risks. Analysts need to account for all factors before making estimates and recommendations.

In the second research objective, statistical estimation techniques were applied to assess the stock prediction accuracy and performance. Vector-auto-regression (VAR) model gives the best predictive performance measured through the model fitness, i.e., R-squared. Though the model produces accurate forecasts, results can be improved with high-frequency data provided to the forecaster/analyst. As a recommendation, the study hints that internal firm operations acquire and utilize high-frequency stock data. The mainframe Enterprise Resource Planning (ERP) framework would optimize other functions reflecting stock performance on a real-time basis.

The statistical models have a limitation of producing estimates with lower confidence intervals. Accuracy issues can be overcome using more granular and time-bound, for example, one week- 2 week ahead predictions using input data.

The third objective implements various artificial neural network architectures that overcome limitations in statistical models in accuracy, performance etc. From the analysis through Automatic Linear modeling, the linearity of market data is observed higher in the BSE index than NSE. Hence, the time-series-based statistical models' prediction performance, i.e., ARMAX, gives better accuracy. The artificial neural network models like MLP (Multilayer perceptron model) and NARX (Non-linear autoregressive models) give higher prediction performance for the NSE prediction tasks. Still, architectures like RBF (Radial basis function) offered deficient performance, which contradicts results compared to Komo *et al*. 1994 and Sheelapriya and Murugesan, 2017. The predictive results could have been affected by the internal algorithm of weights correction, but an augmentation for this is out of scope. One of the technical inferences from the findings is that while radial basis functions generate good approximation in engineering applications, their suitability to model financial markets is limited. The decision support models need to implement either feedback correcting system like MLP and a better algorithm to fine-tune predictive accuracy and residual values. Research also paved the way for marginal improvement in convergence rate and precision level earlier reported for stock market prediction using the NARX model by Alkhoshi and Belkasim, 2018. Also, since Gandhmal and Kumar, 2021 only applied the NARX model in the telecom and manufacturing sector, focusing on two firms' predictions, the current study is more applicable in the financial industry.

The last objective explored possibilities of enhancing artificial neural network prediction model performance by assessing various technologies surveying state-of-art prediction and information systems support tools. Predictive performance after the inclusion of qualitative data like stock price search interest as a predictor has been mixed. Such finding aligns with both developed markets studies and emerging economies (Nardo *et al*. 2016; Pyo *et al*. 2017). Hence the tradeoff of optimizing neural network prediction model using Google trends data produces lesser robust results.

Other technological artifacts and their role in enhancing the predictive capabilities of DSS are explored as an alternative solution. There have been preliminary results that point to blockchain-supported technologies in aiding the decision support tools. An optimal prediction and decision support model is feasible to utilize high-frequency data or similar advanced for market prediction. The study obtained insights into the need for information system modules that integrate internal and external processes through this objective. The stock market prediction and operations are just tasks incorporated in newer models that provide valuable suggestions given input parameters on the market and firm variables. In contrast, most of these aspects will require a significant shift through Business Process re-engineering (BPR). Current findings support the view for fostering a blockchain through a global inter-payment network between banks, expanding firm competency.

5.2. CONTRIBUTIONS TO BODY OF KNOWLEDGE

The research study has added insights to the body of knowledge in multiple areas. In terms of theoretical interpretation, the random walk hypothesis supportive studies in the Indian context have been disproven for short-term predictions in the banking sector. Using several statistical techniques and neural network models, the predictive relevance of these points implies weak random walk theory in Indian capital markets of bank firms. While the efficacy of different predictive models was tested, the results also suggest using high-frequency data in market prediction tasks. Currently, in India, high-frequency trading is still at a nascent stage. The developed economies like the USA, UK etc., already employ algorithmic trading that changes the unique features followed in traditional markets. Another main area that has gained insights pertains to implementing a predictive model for the BSE/NSE banking index. Current research gathers few contradicting results concerning prior literature, particularly in radial basis function models. While qualitative factors like search price interest can be added, its lesser accuracy affects the decision support perspective. The technology and processes as far as India are concerned started during 2010. In later years, whitepapers were published from NSE and BSE. A few factors support industry practice based on technology up-gradation for emerging markets (Kauffman *et al*. 2015).

Even though this was the case, high-frequency trading still has no accepted regulatory or legal definition, among other issues. Only through good policy-making, technological interface overcoming digital divide and infrastructure can such technologies gain traction. The development of expert-level decision support systems requires the inclusion of a variety of data. In the current research, the predictive model utilized only available secondary data extracted from standard databases of stock exchanges. Even if data sources are implemented, it requires considerable experience and efficiency from analysts and other stakeholders.

5.3. IMPLICATIONS AND LIMITATIONS

5.3.1. Technical implications

The research study shows that specific decision support system models need to be designed using predictive algorithms. Some of them are backpropagation and multilayer feedforward networks specific to real-time input market data. Periodic iterations only can ensure model fitness and accuracy for analysts' recommendations, policy etc. New industry-led consortiums led blockchain systems such as R3, and FinTech 2.0 strategy formation is essential in developing economies. Instead of replicating corporate strategies such as Google, Amazon, Apple or Facebook, digital banking systems need to be redesigned for new modalities. The Paytm of India is an example wherein a localized market-specific operational strategy was adopted. Most of these firms are developing in-house payments ex: Apple Pay, Near Field Technology (NFC) enabled Google Pay, contactless payments, or facial recognition. Therefore, a paradigm shift to the evolution of banking models on the services and operations front is key to staying relevant. The synergy of information systems for banking management with a real-time interbank network worldwide is the future. By monitoring global risks, functionality is augmented. Processes include interactive and business intelligence tools for mobile banking, stock market predictions, credit evaluations, risk management, etc., for employees, investors and customers.

5.3.2. Managerial implications

Already studies have shown that various sectors of the Indian capital market are affected by economic fundamentals that combine sectoral indices in the long run. BANKEX is the dominant driver for integration and the spillover of information (Vardhan *et al*. 2015). The banking firms competing in the business world must optimize both internal and external resources. Specifically, the aspects of stock exchange operations and financial disclosure significantly impact future market performance. Research in developed markets has explored the relationship between user-generated content through social platforms and market performance of stocks, so that appropriately managerial actions are tweaked (Tirunillai and Tellis, 2012). The standardization in transactions, privacy/security infrastructure and resolution mechanism by large banking institutions can influence smaller and non-banking public and private blockchain entities. New corporate social media and customer relations management will hence be solid catalysts for operations strategy. Psychological aspects like risk-taking, trend forecast ability of professional traders from investors and analyst perspectives need to be given due importance before planning market operations. (Häusler *et al*. 2018; Glaser *et al*. 2007)

5.3.3. Limitations of research

The stock market operating in any economy is a highly complicated system for modeling or prediction purposes due to many variables. Especially, the problem becomes compounded when social dynamics are involved in the process (Hofman *et al*. 2017). As discussed earlier in the study, these can range from macro and micro indicators, i.e., purely financial, economic, political factors, investor behavior, attitude etc. (Eickhoff and Muntermann, 2016; Jacaruso, 2018; Setiani *et al*. 2021). In the current study, primarily historical market data is utilized in predictive model building. But in practice, analysts use various additional parameters to forecast market trends that contribute to public information (Altınkılıç *et al*. 2013; Singh and Khushi, 2021).

5.4. FUTURE DIRECTIONS AND SCOPE

There are many directions to build on existing results presented in research or extend to further empirical work. These include methodological extensions or model enhancement through additional parameters. Such aspects are outlined, discussing the recent works in this subsection. The emergence of big data frameworks in the stock market has been seen in the last decade. The primary reason for adopting such methods can be increased computational capacity and the proliferation of information technology-based tools worldwide. Some of the studies on this line of research used methods like sentiment analysis, machine learning. Finally, they suggested that Google trends data inclusion is future scope. Also, learning vector quantization gave better results over other compared algorithms.

These results indicate that stock prediction models supported by big data computing frameworks have the potential to forecast and estimate trends much effectively than existing models. Usage of multiple information sources can lead to better prediction performance and accuracy (Wang *et al*. 2019; Weng *et al*. 2017). In these lines, a recent study investigated information and flow. Here, several factors can be controlled that determine trading decisions, leading to stock market activity. Some of these are social interactions, interpersonal communications, accounting, attitude and personality, internet news, social media, and specialist media. Each of these factor parameters in prediction models needs further research. The markets have shown positive or negative impacts from the diffusion of news events into media outlets from the earliest literature sources. Several studies have been explored these phenomena to quantify the effects of news on markets. On this aspect, a study using LSTM (Long Short Term Memory) is an advanced neural network model. The granger causality test that banking stocks from these studies suggest that technology-related news heavily affects the market. Due to the explosion of social media as a significant force globally, many researchers have analyzed its role in financial markets (Atkins *et al*. 2018; Li *et al*. 2020). Extant literature finds that using social media data exists predictive relationships on stock volume and volatility.

On similar lines, Twitter reactions based on event study methodology can offer rich insights (Bollen *et al*. 2011). Some studies indicate that the volume and sentiment of industry tweets can predict share price movement across firms and industries. Results from analysis of microblogs of tweet data opining that market inefficiencies can be exploited in results. Effects of the management disclosure on information sharing forums also call for empirical work on these areas. Similar work has been conducted like studies that assess the process through which information disclosure by official sources affects investors and the market environment (Xu *et al*. 2013). Additionally, methodological extensions exist, such as deep learning; genetic algorithms for market prediction are still open (Chong *et al*. 2017; Thakkar and Chaudhari, 2021b). Even the marketing decisions for the firms can be extended to prediction models as decision support applications.

To further optimize a prediction model, particle swarm optimization (PSO) algorithms are useful (Senapati *et al*. 2018). The mandatory firm disclosures through regulatory forums utilized text mining techniques for stock forecasting offers insights into long-term prediction models as extended research (De Fortuny *et al*. 2014; Feuerriegel and Gordon, 2018; Gottschlich and Hinz, 2014). The incremental and unique contribution from research effectively integrates previously untested statistical models and machine learning methods, namely neural networks for stock market index prediction. By doing so, it also integrates secondary market data that combines fundamental and technical factors for forecasts. The efficacy of integrating qualitative trend data from the internet into the decision-making framework is also investigated following prior literature. These results can be successfully be used in better market trading systems, stakeholders or businesses. Pursuing empirical research in the above directions holds vast potential to improve and build better predictive models than the existing state-of-art. It is of value when such studies are being placed in the context of India so that market prediction and support functionalities can be fine-tuned to local scenarios. Current findings from research literature strongly encourage the design of better management information systems. Finally, the study intends to provide a strategic roadmap in implementing best practices, design long-term policies that foster growth and co-create value to firms supported by empirical research.

APPENDICES

TABLE 4.3.7: PREDICTIVE ACCURACY SUMMARY

Forecast method	**Prediction performance**		**Independent variable importance (normalized 0-least important to 5-very important or in % values)**	
	BSE Bankex	**NSE Nifty Bank**	**BSE Bankex**	**NSE Nifty Bank**
Automatic Linear Modelling (ALM)	AIC= 218.660 Accuracy= 91.5%	AIC= 214.58 Accuracy= 83.3%	High index (0.85) Div. Yield (0.25)	High index (0.65) P/B ratio (0.19) P/E ratio (0.1) Beta (0.05)
ARMAX	Stationary R^2 = 0.976	Stationary R^2 = 0.955	N/A	N/A
MLP	Sum of squares errors= 0.005 Relative error = 0.001	Sum of squares errors = 0.002 Relative error = 0.000	Div. Yield = 100% High index = 90.02% Beta = 87.5% P/B ratio = 81.2% P/E ratio = 75.5%	P/B ratio = 100% P/E ratio = 88.3% Div. yield = 78.6% High index = 61.2 Beta = 38.2%

RBF	Sum of squares errors = 1.652 Relative error = 0.254	Sum of squares errors = 3.669 Relative error = 0.667	Beta = 100% Div. yield = 9.1% All other factors = 9.4	Beta = 100% All other factors = 17.7%
VAR	Adjusted R^2= 0.97 BIC = 17.12 Durbin-Watson = 1.97	Adjusted R^2 = 0.99 BIC = 13.18 Durbin-Watson = 1.98	N/A	N/A
NARX	Overall model R^2 = 0.99823	Overall model R^2 = 0.9753	N/A	N/A

Authors' calculation

TABLE 4.3.9: GLOBAL STOCK INDICES CORRELATIONS

			Stock price (World wide)	**BSE Bombay**	**STI Singapore**	**DJI USA**	**TSX Canada**	**HSI Hongkong**
Spearman's Rho	Stock price (World wide)	Correlation coefficient Sig. (2-tailed) N	1.000 . 60	.792** .000 60	.180 .169 60	.874** .000 60	.622* .000 60	.180 .169 60
	BSE Bombay	Correlation coefficient Sig. (2-tailed) N	.792* 0.000 60	1.000 . 60	.605** .000 60	.916** .000 60	.798** .000 60	.605** .000 60
	STI Singapore	Correlation coefficient Sig. (2-tailed) N	.180 .169 60	.605** .000 60	1.000 . 60	.474** .000 60	.662** .000 60	1.000** . 60
	DJIA USA	Correlation coefficient	.874** .000 60	.916** .000 60	.474** .000 60	1.000 . 60	.817** .000 60	.474** .000 60

		Sig. (2-tailed) N						
	TSX Canada	Correlation coefficient Sig. (2-tailed) N	.622** .000 60	.798** .000 60	.662** .000 60	.817** .000 60	1.000 . 60	.662** .000 60
	HSI Hongkong	Correlation coefficient Sig. (2-tailed) N	.180 .169 60	.605** .000 60	1.000** . 60	.474** .000 60	.662** .000 60	1.000 . 60

Authors' calculation

*** Correlation is significant at the 0.01 level (2- tailed)*

TABLE 4.4.1: TOP IMPACT MAKING 5 STUDIES (SOURCE: SCOPUS)

Author/ Year of Study	**Objective & Methods**	**Findings and Limitations**	**Central theme and Suggestions.**	**Total citations[1] (TC)**
Guo and Liang (2016)	To assess the blockchain for the banking industry in the Chinese context. Uses primary data and industry reports.	Blockchain + banks (FinTech 2.0) has superior customer experience, efficiency, cost and security.	Theme: Banking. Shift to R3, a consortium-led blockchain. Payment clearing systems, bank credit information systems are vital.	91
Zhao *et al.* (2016)	*Editorial note*. Introduction to 7 academic papers in a special edition.	Multidisciplinary nature of blockchain research. Wos and SSRN articles grew from 0 in 2014 to 11 and 79, respectively, by 2016.	Theme: Innovation. Challenges remain on theoretical issues of societal impact, smart contracts, security etc.	57
Kiviat (2015)	To integrate research from diverse domains and educate the legal community, help practitioners. Primary and secondary data, legal proceedings analysis.	Regulatory patchworks are underway in the USA, but virtual currencies pose uncertainty.	Theme: Legal. Policymakers must be cautious and have precision in tailoring the scope of regulation.	44

Nguyen (2016)	To study the role of blockchain in sustainable outcomes investigating in the finance industry.	Limits competition as a network is shared. Payment risks, technology cost issues.	Theme: Sustainability Further study on crowdsourcing, payment system. Banks face market pressure.	33
Larios-Hernández (2017)	To explain the lack of a formal bank account. Leverage Blockchain-based financial inclusion. Fuzzy-set Qualitative Comparative Analysis (fsQCA)	Variety of supply-and-demand related causal factors. Informal peer-to-peer practice is customary.	Theme: Entrepreneurship/ financial inclusion. The business logic of existing financial institutions hinders the solutions for the unbanked.	27

[1] *Citations data from Scopus database as of January 2020*

TABLE 4.4.2: TOP IMPACT MAKING 5 STUDIES (SOURCE: WEB OF SCIENCE)

Author/ Year of Study	**Objective & Methods**	**Findings and Limitations**	**Central theme and Suggestions.**	**Total citations[2] (TC)**
Gomber *et al*. 2018	To review the technology innovation, process disruption, & services transformations. Quantitative & qualitative data.	Assessing the factors outlined in the study and discussed needs to be repeated.	Theme: Information systems. Scope for developing research by interdisciplinary sources, designs, theory and methodologies.	28
Eyal (2017)	To gain an overview of protocols, distributed-ledger technology (DLT). Exploration of boundaries of blockchain beyond bitcoin.	Requirements of blockchains for cryptocurrencies with FinTech vary drastically— Ex: transaction throughput, security & privacy. A clear distinction between FinTech and cryptocurrencies.	Theme: Financial applications. Direct and effective collaboration is required between the FinTech industry and the blockchain scientific engineering community.	25
Anjum *et al*. 2017	To compare the types and performance of blockchain architectures.	Platforms for specific sectors and application domains are emerging. Current models have limitations in scalability, flexibility, & governance.	Theme: Standards. Standardization activity is required as technology enablers and for interoperability.	23

Treleaven *et al*. 2017	*Editorial note*. On special issue covering journal research articles.	As blockchain technology evolves, it becomes disruptive for other technologies such as big data, the IoT, intelligent assistants, and autonomous vehicles with opportunities and unintended social consequences.	Multi-themed. Smart contracts could become the management framework for many private records in the future.	21
Ziegeldorf *et al*. 2018	To study *CoinParty*, an efficient decentralized mixing service for users to reestablish financial privacy.	The prototype implementation scales to many users achieve data anonymity higher level than earlier models.	Theme: Data privacy. Third-party independent deployment is feasible for organizations.	15

[2] *Citations data from Web of Science database as of January 2020*

REFERENCES

Abreu, M. and Mendes, V. (2012). Information, overconfidence and trading: Do the sources of information matter?. *Journal of Economic Psychology*, *33*(4), 868-881.

Accenture Inc. 2016, White paper. Blockchain Technology. Retrieved from https://www.accenture.com/_acnmedia/pdf-35/accenture-blockchain-how-banks-building-real-time-global-payment-network.pdf . Accessed on 20-02-2020.

Ahmed, M., Sriram, A. and Singh, S. (2019). Short Term Firm-Specific Stock Forecasting with BDI Framework. *Computational Economics*, 1-34.

Agarwal, A., Leung, A. C. M., Konana, P. and Kumar, A. (2017). Cosearch attention and stock return predictability in supply chains. *Information Systems Research*, *28*(2), 265-288.

Agarwal, S., Kumar, S. and Goel, U. (2019). Stock market response to information diffusion through internet sources: A literature review. *International Journal of Information Management*, *45*, 118-131.

Aithal, P. K., Dinesh, A. U. and Geetha, M. (2019). Identifying Significant Macroeconomic Indicators for Indian Stock Markets. *IEEE Access*, *7*, 143829-143840.

Alkhoshi, E. and Belkasim, S. (2018, September). Stable Stock Market Prediction Using NARX Algorithm. In *Proceedings of the 2018 International Conference on Computing and Big Data* (pp. 62-66). ACM.

Allen, H. and Taylor, M. P. (1990). Charts, noise and fundamentals in the London foreign exchange market. *The Economic Journal*, *100*(400), 49-59.

Altınkılıç, O., Balashov, V. S. and Hansen, R. S. (2013). Are Analysts' Forecasts Informative to the General Public? *Management Science*, *59*(11), 2550–2565.

Andreassen, P. B. (1987). On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction. *Journal of Personality and Social Psychology*, *53*(3), 490.

Andreassen, P. B. (1988). Explaining the price-volume relationship: The difference between price changes and changing prices. *Organizational Behavior and Human Decision Processes*, *41*(3), 371-389.

Anjum, A., Sporny, M. and Sill, A. (2017). Blockchain standards for compliance and trust. *IEEE Cloud Computing*, *4*(4), 84-90.

Arasu, B. S., Jeevananthan, M., Thamaraiselvan, N. and Janarthanan, B. (2014). Performances of data mining techniques in forecasting stock index–evidence from India and US. *Journal of the National Science Foundation of Sri Lanka*, *42*(2).

Ariyo, A. A., Adewumi, A. O. and Ayo, C. K. (2014, March). Stock price prediction using the ARIMA model. In *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation* (pp. 106-112). IEEE.

Armano, G., Marchesi, M. and Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences*, *170*(1), 3-33.

Aspara, J. and Tikkanen, H. (2010). Consumers' stock preferences beyond expected financial returns: The influence of product and brand evaluations. *International Journal of Bank Marketing*, *28*(3), 193-221.

Atkins, A., Niranjan, M. and Gerding, E. (2018). Financial news predicts stock market volatility better than close price. *The Journal of Finance and Data Science*, *4*(2), 120-137.

Atsalakis, G. S. and Valavanis, K. P. (2009). Surveying stock market forecasting techniques–Part II: Soft computing methods. *Expert Systems with Applications*, *36*(3), 5932-5941.

Attigeri, G. V., MM, M. P., Pai, R. M. and Nayak, A. (2015, November). Stock market prediction: A big data approach. In *TENCON 2015-2015 IEEE Region 10 Conference* (pp. 1-5). IEEE.

Babu, C. N. and Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. *Applied Soft Computing*, *23*, 27-38.

Baker, R. and Belgorodskiy, A. (2007). Predicting stock returns and assessing prediction performance. *IMA Journal of Management Mathematics*, *18*(4), 413-433.

Balaji, A.J., Ram, D.H. and Nair, B.B. (2018). Applicability of deep learning models for stock price forecasting an empirical study on Bankex data. *Procedia Computer Science*, *143*, 947-953.

Barrett, J. F. and Wright, D. J. (1974). The random nature of stock-market prices. *Operations Research*, *22*(1), 175-177.

Birau, R., Trivedi, J., and Antonescu, M. (2015). Modeling S&P Bombay stock exchange BANKEX index volatility patterns using GARCH model. *Procedia Economics and Finance*, *32*(1), 520-525.

Bollen, J., Mao, H. and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1-8.

Bonanno, G., Vandewalle, N. and Mantegna, R. N. (2000). Taxonomy of stock market indices. *Physical Review E*, *62*(6), R7615.

Borovkova, S. and Tsiamas, I. (2019). An ensemble of LSTM neural networks for high-frequency stock market classification. *Journal of Forecasting*, *38*(6), 600-619.

Braun, H. and Chandler, J. S. (1987). Predicting stock market behavior through rule induction: an application of the learning-from-example approach. *Decision Sciences*, *18*(3), 415-429.

Brookmire, J. H. (1913). Methods of business forecasting based on fundamental statistics. *The American Economic Review*, *3*(1), 43-58.

Cengiz, C. B. and Herwartz, H. (2011). Modeling stock index returns by means of partial least-squares methods: An out-of-sample analysis for three stock markets. *Applied Stochastic Models in Business and Industry*, *27*(3), 253-266.

Chang, V. and Ramachandran, M. (2017). Financial modeling and prediction as a service. *Journal of Grid Computing*, *15*(2), 177-195.

Charumathi, B. and Suraj, E. S. (2014). Refining ohlson model for valuing bank stocks-an artificial neural network approach. *Independent Business Review*, *7*(2), 32.

Chen, M. Y., Liao, C. H. and Hsieh, R. P. (2019a). Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach. *Computers in Human Behavior*, *101*, 402-408.

Chen, T. H., Ho, R. J. and Liu, Y. W. (2019b). Investor personality predicts investment performance? A statistics and machine learning model investigation. *Computers in Human Behavior*, *101*, 409-416.

Chen, T. L., Cheng, C. H. and Liu, J. W. (2019c). A Causal Time-Series Model Based on Multilayer Perceptron Regression for Forecasting Taiwan Stock Index. *International Journal of Information Technology & Decision Making (IJITDM)*, *18*(06), 1967-1987.

Cheng, J. H., Chen, H. P. and Lin, Y. M. (2010). A hybrid forecast marketing timing model based on probabilistic neural network, rough set and C4. 5. *Expert systems with Applications*, *37*(3), 1814-1820.

Chenoweth, T., Obradovic, Z. and Lee, S. S. (1996). Embedding technical analysis into neural network based trading systems. *Applied Artificial Intelligence*, *10*(6), 523-542.

Chiang, W. C., Enke, D., Wu, T. and Wang, R. (2016). An adaptive stock index trading decision support system. *Expert Systems with Applications*, *59*, 195-207. DOI:

Choi, H. and Varian, H. (2012). Predicting the present with Google Trends. *Economic record*, *88*, 2-9.

Cho, V. (2010). MISMIS–A comprehensive system for stock market investment. *Knowledge-Based Systems*, *23*(6), 626-633.

Chong, E., Han, C. and Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, *83*, 187-205.

Chottiner, S. (1972). Stock Market Research Methodology: A Case for the Systems Approach. *Decision Sciences*, *3*(2), 45-53.

Coakley, J. R. and Brown, C. E. (2000). Artificial neural networks in accounting and finance: Modeling issues. *Intelligent Systems in Accounting, Finance and Management*, *9*(2), 119-144.

Cooper, D. and Schindler, P. Business Research Methods 11/e. McGraw Hill

Cooper, M. J., Jackson III, W. E. and Patterson, G. A. (2003). Evidence of predictability in the cross-section of bank stock returns. *Journal of Banking & Finance*, *27*(5), 817-850.

Cootner, P. H. (1967). *The random character of stock market prices*. MIT press.

Cowles 3rd, A. (1933). Can stock market forecasters forecast?. *Econometrica: Journal of the Econometric Society*, 309-324.

Davamanirajan, P., Kauffman, R. J., Kriebel, C. H. and Mukhopadhyay, T. (2006). Systems design, process performance, and economic outcomes in international banking. *Journal of Management Information Systems*, *23*(2), 65-90.

De Fortuny, E. J., De Smedt, T., Martens, D. and Daelemans, W. (2014). Evaluating and understanding text-based stock price prediction models. *Information Processing & Management*, *50*(2), 426-441.

De Oliveira, F. A., Nobre, C. N. and Zarate, L. E. (2013). Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index–Case study of PETR4, Petrobras, Brazil. *Expert Systems with Applications*, *40*(18), 7596-7606.

Dewan, S. and Mendelson, H. (1998). Information technology and time-based competition in financial markets. *Management Science*, *44*(5), 595-609.

Dimpfl, T. and Jank, S. (2016). Can internet search queries help to predict stock market volatility?. *European Financial Management*, *22*(2), 171-192.

Dincer, H. (2015). Profit-based stock selection approach in banking sector using Fuzzy AHP and MOORA method. *Global Business and Economics Research Journal*, *4*(2), 1-26.

Ding, X., Zhang, Y., Liu, T. and Duan, J. (2015, June). Deep learning for event-driven stock prediction. In *Twenty-fourth international joint conference on artificial intelligence*.

Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, *72*(1), 1-16.

Dutta, G., Jha, P., Laha, A. K. and Mohan, N. (2006). Artificial neural network models for forecasting stock price index in the Bombay stock exchange. *Journal of Emerging Market Finance*, *5*(3), 283-295.

Earley, P. C., Connolly, T. and Ekegren, G. (1989). Goals, strategy development, and task performance: Some limits on the efficacy of goal setting. *Journal of Applied Psychology*, *74*(1), 24.

East, R. (1993). Investment decisions and the theory of planned behaviour. *Journal of Economic Psychology*, *14*(2), 337-375.

Eickhoff, M. and Muntermann, J. (2016). Stock analysts vs. the crowd: Mutual prediction and the drivers of crowd wisdom. *Information & Management*, *53*(7), 835-845.

Eyal, I. (2017). Blockchain technology: Transforming libertarian cryptocurrency dreams to finance and banking realities. *Computer*, *50*(9), 38-49.

Fadlalla, A. and Amani, F. (2014). Predicting next trading day closing price of Qatar exchange index using technical indicators and artificial neural networks. *Intelligent Systems in Accounting, Finance and Management*, *21*(4), 209-223.

Favero, C. A., Gozluklu, A. E. and Tamoni, A. (2011). Demographic trends, the divid end-price ratio, and the predictability of long-run stock market returns. *Journal of Fin ancial and Quantitative Analysis*, 1493-1520.

Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L. and Muller, P. A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, *33*(4), 917-963.

Felsen, J. (1975). Learning pattern recognition techniques applied to stock market forecasting. *IEEE Transactions on Systems, Man, and Cybernetics*, (6), 583-594.

Feuerriegel, S. and Gordon, J. (2018). Long term stock index forecasting based on text mining of regulatory disclosures. *Decision Support Systems*, *112*, 88-97.

French, J. W. (1972). Application of T-technique factor analysis to the stock market. *Multivariate Behavioral Research*, *7*(3), 279-286.

Gabor, D. and Brooks, S. (2017). The digital revolution in financial inclusion: international development in the fintech era. *New Political Economy*, *22*(4), 423-436.

Gandhmal, D. P. and Kumar, K. (2021). Wrapper-Enabled Feature Selection and CPLM-Based NARX Model for Stock Market Prediction. *The Computer Journal*, *64*(2), 169-184.

Gao, T., Li, X., Chai, Y. and Tang, Y. (2016, August). Deep learning with stock indicators and two-dimensional principal component analysis for closing price prediction system. In *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)* (pp. 166-169). IEEE.

Glaser, M., Langer, T. and Weber, M. (2007). On the trend recognition and forecasting ability of professional traders. *Decision Analysis*, *4*(4), 176-193.

Gomber, P., Kauffman, R.J., Parker, C. and Weber, B.W., 2018. On the fintech revolution: interpreting the forces of innovation, disruption, and transformation in financial services. *Journal of Management Information Systems*, *35*(1), pp.220-265.

Gomber, P., Koch, J. A. and Siering, M. (2017). Digital Finance and FinTech: Current research and future research directions. *Journal of Business Economics*, *87*(5), 537–580.

Gottschlich, J. and Hinz, O. (2014). A decision support system for stock investment recommendations using collective wisdom. *Decision Support Systems*, *59*, 52-62.

Goyal, A. and Welch, I. (2003). Predicting the equity premium with dividend ratios. *Management Science*, *49*(5), 639-654.

Greenwood, J. and Jovanovic, B., (1999). The information-technology revolution and the stock market. *American Economic Review*, *89*(2), pp.116-122.

Guldiken, O., Tupper, C., Nair, A. and Yu, H. (2017). The impact of media coverage on IPO stock performance. *Journal of Business Research*, *72*, 24-32.

Guo, Y. and Liang, C. (2016). Blockchain application and outlook in the banking industry. *Financial Innovation*, *2*(1), p.24.

Gupta, A. and Xia, C. (2018). A Paradigm shift in banking: unfolding Asia's FinTech adventures. In *Banking and Finance Issues in Emerging Markets*. Emerald Publishing Limited.

Guresen, E., Kayakutlu, G. and Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, *38*(8), 10389-10397.

Haldane, A. G. and May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, *469*(7330), 351.

Hansen, J. (1956, January). Technical market analysis using a computer. In *Proceedings of the 1956 11th ACM national meeting* (pp. 37-40). ACM Digital library.

Häusler, A. N., Kuhnen, C. M., Rudorf, S. and Weber, B. (2018). Preferences and beliefs about financial risk taking mediate the association between anterior insula activation and self-reported real-life stock trading. *Scientific Reports*, *8*(1), 1-13.

Henrique, B. M., Sobreiro, V. A. and Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications*, *124*, 226-251.

Hickman, K. and Petry, G. H. (1990). A comparison of stock price predictions using court accepted formulas, dividend discount, and P/E models. *Financial Management*, 76-87.

Hiransha, M., Gopalakrishnan, E. A., Menon, V. K. and Soman, K. P. (2018). NSE stock market prediction using deep-learning models. *Procedia Computer Science*, *132*, 1351-1362.

Hofman, J. M., Sharma, A. and Watts, D. J. (2017). Prediction and explanation in social systems. *Science*, *355*(6324), 486-488.

Hoontrakul, P. (2018). Asia's digital economy. In *Economic Transformation and Business Opportunities in Asia* (pp. 269-312). Palgrave Macmillan, Cham.

Hu, D., Zhao, J.L., Hua, Z. and Wong, M.C. (2012). Network-based modeling and analysis of systemic risk in banking systems. *MIS Quarterly*, 1269-1291.

Hu, H., Tang, L., Zhang, S. and Wang, H. (2018). Predicting the direction of stock markets using optimized neural networks with Google Trends. *Neurocomputing*, *285*, 188-195.

Hu, Z., Bao, Y., Chiong, R. and Xiong, T. (2017). Profit guided or statistical error guided? A study of stock index forecasting using support vector regression. *Journal of Systems Science and Complexity*, *30*(6), 1425-1442.

Ibrahim, A., Kashef, R., Li, M., Valencia, E. and Huang, E. (2020). Bitcoin Network Mechanics: Forecasting the BTC Closing Price Using Vector Auto-Regression Models Based on Endogenous and Exogenous Feature Variables. *Journal of Risk and Financial Management*, *13*(9), 189.

Jacaruso, L. C. (2018). A method of trend forecasting for financial and geopolitical data: inferring the effects of unknown exogenous variables. *Journal of Big Data*, *5*(1), 1-14.

Jadhav, S., He, H., & Jenkins, K. (2015, November). Prediction of earnings per share for industry. In *2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)* (Vol. 1, pp. 425-432). IEEE.

Jaffee, D. M., Malkiel, B. G. and Quandt, R. E. (1974). Predicting common stock prices: Payoffs and pitfalls. *Journal of Business Research*, *2*(1), 1-16.

Jiang, C., Liang, K., Chen, H. and Ding, Y. (2014). Analyzing market performance via social media: a case study of a banking industry crisis. *Science China Information Sciences*, *57*(5), 1-18.

Jin, Z., Yang, Y. and Liu, Y. (2019). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications*, 1-17.

Junqué de Fortuny, E., Martens, D. and Provost, F. (2013). Predictive modeling with big data: Is bigger really better?. *Big Data*, *1*(4), 215-226.

Kara, Y., Boyacioglu, M. A. and Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert systems with Applications*, *38*(5), 5311-5319.

Kauffman, R. J., Hu, Y. and Ma, D. (2015). Will high-frequency trading practices transform the financial markets in the Asia Pacific Region?. *Financial Innovation*, *1*(1), 4.

Keller, C. and Siegrist, M. (2006). Investing in stocks: The influence of financial risk attitude and values-related money and stock market attitudes. *Journal of Economic Psychology*, *27*(2), 285-303.

Kelotra, A. and Pandey, P. (2020). Stock Market Prediction Using Optimized Deep-ConvLSTM Model. *Big Data*, *8*(1), 5-24.

Khan, M. T. I., Tan, S. H. and Chong, L. L. (2017). Information sources and investing decisions–a path modeling approach. *Managerial Finance*, *43*(8), 928-947.

Khansa, L. and Liginlal, D. (2011). Predicting stock market returns from malicious attacks: A comparative analysis of vector autoregression and time-delayed neural networks. *Decision Support Systems*, *51*(4), 745-759.

Khatua, A. and Khatua, A. (2015). How Stock Market Reacts to Budget Announcement? Through the Lens of Social Media in Indian Context. *Management and Labour Studies*, *40*(3-4), 239-251.

Kidd, J. B. and Morgan, J. R. (1969). A predictive information system for management. *Journal of the Operational Research Society*, *20*(2), 149-170.

Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, *55*(1-2), 307-319.

Kim, S. K. and Park, J. I. (1996). A structural equation modeling approach to generate explanations for induced rules. *Expert Systems with Applications*, *10*(3-4), 403-416.

Kiviat, T. I. (2015). Beyond Bitcoin: Issues in Regulating Blockchain Transactions. *Duke Law Journal*, *65*(3), 569-608.

Komo, D., Chang, C. I. and Ko, H. (1994, April). Neural network technology for stock market index prediction. In *Proceedings of ICSIPNN'94. International Conference on Speech, Image Processing and Neural Networks* (pp. 543-546). IEEE.

Koutroumanidis, T., Ioannou, K. and Zafeiriou, E. (2011). Forecasting bank stock market prices with a hybrid method: the case of Alpha bank. *Journal of Business Economics and Management*, *12*(1), 144-163.

Krolak, P., Berquist, R., Conn, R. and Gilliland, H. (1969). A simulation model for evaluating the effectiveness of various stock market strategies. *Simulation*, *13*(6), 299-305.

Lakshmi, M. (2013). PLS-Neural Network Model for Stock Prices Prediction. *International Journal on Information Sciences & Computing*, 7(2). p. 15-20.

Lam, M. (2004). Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*, *37*(4), 567-581.

Larios-Hernández, G. J. (2017). Blockchain entrepreneurship opportunity in the practices of the unbanked. *Business Horizons*, *60*(6), 865-874.

LeBaron, B., Arthur, W. B. and Palmer, R. (1999). Time series properties of an artificial stock market. *Journal of Economic Dynamics and control*, *23*(9-10), 1487-1516.

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Leigh, W., Purvis, R. and Ragusa, J. M. (2002). Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *Decision Support Systems*, *32*(4), 361-377.

Leonardi, P. M., Bailey, D. E., Diniz, E. H., Sholler, D. and Nardi, B. A. (2016). Multiplex Appropriation in Complex Systems Implementation: The Case of Brazil's Correspondent Banking System. *MIS Quarterly*, *40*(2), 461-473.

Leung, M. T., Daouk, H. and Chen, A. S. (2000). Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting*, *16*(2), 173-190.

Levin, A. E. (1996). Stock selection via nonlinear multi-factor models. *Advances in Neural Information Processing Systems*, 966-972.

Levy, R. A. (1966). Conceptual foundations of technical analysis. *Financial Analysts Journal*, *22*(4), 83-89.

Li, X., Cao, J. and Pan, Z. (2019). Market impact analysis via deep learned architectures. *Neural Computing and Applications*, *31*(10), 5989-6000.

Li, L., Tang, Y., Li, B. and Wu, X. (2004). Artificial neural network based universal time series. *U.S. Patent No. 6,735,580*. Washington, DC: U.S. Patent and Trademark Office.

Li, T., van Dalen, J. and van Rees, P. J. (2018). More than just noise? Examining the information content of stock microblogs on financial markets. *Journal of Information Technology*, *33*(1), 50-69.

Li, Y., Spigt, R. and Swinkels, L. (2017). The impact of FinTech start-ups on incumbent retail banks' share prices. *Financial Innovation*, *3*(1), 26.

Li, X., Wu, P. and Wang, W. (2020). Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong. *Information Processing & Management*, *57*(5), 102212.

Li, Z., Lyu, S., Zhang, H. and Jiang, T. (2021). One Step Ahead: A Framework for Detecting Unexpected Incidents and Predicting the Stock Markets. *IEEE Access*, *9*, 30292-30305.

Lin, M. C., Lee, A. J., Kao, R. T. and Chen, K. T. (2008). Stock price movement prediction using representative prototypes of financial reports. *ACM Transactions on Management Information Systems (TMIS)*, *2*(3), 1-18.

Lin, Q. (2018). Technical analysis and stock return predictability: An aligned approach. *Journal of Financial Markets*, *38*, 103-123.

Liu, X. and Yu, T. (2018). An automatic pattern recognition value system with listed banks based on blockchain. In *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (pp. 1850-1854).

Lo, A.W. (2004). The adaptive markets hypothesis. *The Journal of Portfolio Management*, *30*(5), 15-29.

Lo, A. W., Mamaysky, H. and Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, *55*(4), 1705-1765.

Loh, L. and Ong, Y. S. (1998). The adoption of Internet-based stock trading: a conceptual framework and empirical results. *Journal of Information Technology*, *13*(2), 81-94.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417.

Maji, G., Sen, S. and Sarkar, A. (2017, June). Share Market Sectoral Indices Movement Forecast with Lagged Correlation and Association Rule Mining. In *IFIP International Conference on Computer Information Systems and Industrial Management*. 327-340. Springer, Cham.

Malik, F., Wang, F. and Naseem, M. A. (2017). Econometric estimation of banking stocks. *The Journal of Developing Areas*, *51*(4), 207-237.

Meenakshi, S. S. and Lakshmi, M. (2013). PLS-Neural Network Model for Stock Prices Prediction. *International Journal on Information Sciences and Computing*, *7*(2).

Miraz, M. H. and Donald, D. C. (2018, August). Application of blockchain in booking and registration systems of securities exchanges. In *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)* (pp. 35-40). IEEE.

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E. and Preis, T. (2013). Quantifying Wikipedia usage patterns before stock market moves. *Scientific reports*, *3*(1), 1-5.

Moghaddam, A. H., Moghaddam, M. H. and Esfandyari, M. (2016). Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, *21*(41), 89-93.

Mohapatra, S. and Misra, A. K. (2019). Cross-sectional returns predictability for emerging market banks: A study on Indian banking system. *Cogent Economics & Finance*, *7*(1).

Moro, S., Cortez, P. and Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent Dirichlet allocation. *Expert Systems with Applications*, *42*(3), 1314-1324.

Nagle, F. (2013). Stock market prediction via social media: The importance of competitors. In *Academy of Management Proceedings*, Vol. 2013, No. 1, p. 17557. Briarcliff, Academy of Management.

Nardo, M., Petracco-Giudici, M. and Naltsidis, M. (2016). Walking down wall street with a tablet: A survey of stock market predictions using the web. *Journal of Economic Surveys*, *30*(2), 356-369.

Nayak, R. K., Mishra, D. and Rath, A. K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, *35*, 670-680.

Nayak, S. C. and Misra, B. B. (2018). Estimating stock closing indices using a GA-weighted condensed polynomial neural network. *Financial Innovation*, *4*(1), 21.

Neely, C. J., Rapach, D. E., Tu, J. and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators. *Management Science*, *60*(7), 1772-1791.

Nguyen, Q. K. (2016, November). Blockchain-A Financial Technology for Future Sustainable Development. In *2016 3rd International Conference on Green Technology and Sustainable Development (GTSD)* (pp. 51-54).

Nofer, M. (2015). Are Crowds on the Internet Wiser than Experts?–The Case of a Stock Prediction Community. In *The Value of Social Media for Predicting Stock Returns* (pp. 27-61). Springer Vieweg, Wiesbaden.

Nti, I. K., Adekoya, A. F. and Weyori, B. A. (2019). A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*, 1-51.

Nti, I. K., Adekoya, A. F. and Weyori, B. A. (2021). A novel multi-source information-fusion predictive framework based on deep neural networks for accuracy enhancement in stock market prediction. *Journal of Big Data*, *8*(1), 1-28.

Olson, D. and Mossman, C. (2003). Neural network forecasts of Canadian stock returns using accounting ratios. *International Journal of Forecasting*, *19*(3), 453-465.

Ouyang, H., Zhang, X. and Yan, H. (2019). Index tracking based on deep neural network. *Cognitive Systems Research*, *57*, 107-114.

Oztekin, A., Kizilaslan, R., Freund, S. and Iseri, A. (2016). A data analytic approach to forecasting daily stock returns in an emerging market. *European Journal of Operational Research*, *253*(3), 697-710.

Palsson, M.S., Gu, M., Ho, J., Wiseman, H.M. and Pryde, G.J. (2017). Experimentally modeling stochastic processes with less memory by the use of a quantum processor. *Science Advances*, *3*(2), p.e1601302.

Paluch, M. and Jackowska-Strumiłło, L. (2018). Hybrid models combining technical and fractal analysis with ANN for short-term prediction of close values on the warsaw stock exchange. *Applied Sciences*, *8*(12), 2473.

Panda, C. and Narasimhan, V. (2006). Predicting stock returns: an experiment of the artificial neural network in Indian stock market. *South Asia Economic Journal*, *7*(2), 205-218.

Pascual-Ezama, D., Scandroglio, B and Gil-Gómez, B. (2014). Can we predict individual investors' behavior in stock markets? A psychological approach. *Universitas Psychologica, 13(1),* 25-35.

Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*, *42*(1), 259-268.

Patra, B. and Padhi, P. (2015). Backtesting of value at risk methodology: analysis of banking shares in India. *Margin: The Journal of Applied Economic Research*, *9*(3), 254-277.

Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, *62*(3), 54-60.

Perlin, M. S., Caldeira, J. F., Santos, A. A. and Pontuschka, M. (2017). Can we predict the financial markets based on Google's search queries?. *Journal of Forecasting*, *36*(4), 454-467.

Pointer, L. V. and Khoi, P. D. (2019). Predictors of Return on Assets and Return on Equity for Banking and Insurance Companies on Vietnam Stock Exchange. *Entrepreneurial Business and Economics Review*, *7*(4), 185-198.

Pop, C., Pop, C., Marcel, A., Vesa, A., Petrican, T., Cioara, T., Anghel, I. and Salomie, I. (2018). Decentralizing the stock exchange using blockchain an ethereum-based implementation of the Bucharest Stock Exchange. In *2018 IEEE 14th International Conf. on Intelligent Computer Communication and Processing (ICCP)* (pp. 459-66).

Pradhan, R. P., Arvin, M. B., Hall, J. H. and Bahmani, S. (2014). Causal nexus between economic growth, banking sector development, stock market development, and other macroeconomic variables: The case of ASEAN countries. *Review of Financial Economics*, *23*(4), 155-173.

Preis, T., Moat, H. S and Stanley, H. E. (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. *Scientific Reports*, 3.

Pyo, S., Lee, J., Cha, M., & Jang, H. (2017). Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets. *PloS one*, *12*(11), e0188107.

Saud, A. S. and Shakya, S. (2020). Analysis of look back period for stock price prediction with RNN variants: A case study on banking sector of NEPSE. *Procedia Computer Science*, *167*, 788-798.

Qiu, M. and Song, Y. (2016). Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model. *PloS one*, *11*(5), e0155133.

Ramesh, V. P., Baskaran, P., Krishnamoorthy, A., Damodaran, D. and Sadasivam, P. (2019). Back propagation neural network based big data analytics for a stock market challenge. *Communications in Statistics-Theory and Methods*, *48*(14), 3622-3642.

Refenes, A. N., Zapranis, A. and Francis, G. (1994). Stock performance modeling using neural networks: a comparative study with regression models. *Neural networks*, *7*(2), 375-388.

Rihani, V. and Garg, S. K. (2006). Neural networks for the prediction of stock market. *IETE Technical Review*, *23*(2), 113-117.

Majhi, R., Panda, G. and Sahoo, G. (2009). Development and performance evaluation of FLANN based model for forecasting of stock markets. *Expert systems with Applications*, *36*(3), 6800-6808.

Rout, M., Majhi, B., Mohapatra, U. M. and Mahapatra, R. (2012, December). Stock indices prediction using radial basis function neural network. In *International Conference on Swarm, Evolutionary, and Memetic Computing* (pp. 285-293). Springer, Berlin, Heidelberg.

Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536.

Samaras, G. D., Matsatsinis, N. F. and Zopounidis, C. (2008). A multicriteria DSS for stock evaluation using fundamental analysis. *European Journal of Operational Research*, *187*(3), 1380-14010.

Samarawickrama, A. J. P. and Fernando, T. G. I. (2017, December). A recurrent neural network approach in predicting daily stock prices an application to the Sri Lankan stock market. In *2017 IEEE International Conference on Industrial and Information Systems (ICIIS)* (pp. 1-6).

Schinckus, C. (2018). An essay on financial information in the era of computerization. *Journal of Information Technology*, *33*(1), 9-18.

Schultz, H. (1925). Forecasting security prices. *Journal of the American Statistical Association*, *20* (150), 244-249.

Schumaker, R. P. and Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Transactions on Information Systems (TOIS)*, *27*(2), 1-19.

Selvamuthu, D., Kumar, V. and Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, *5*(1), 16.

Senapati, M. R., Das, S. and Mishra, S. (2018). A Novel Model for Stock Price Prediction Using Hybrid Neural Network. *Journal of The Institution of Engineers (India): Series B*, *99*(6), 555-563.

Setiani, I., Tentua, M. N. and Oyama, S. (2021, March). Prediction of Banking Stock Prices Using Naïve Bayes Method. In *Journal of Physics: Conference Series* (Vol. 1823, No. 1, p. 012059). IOP Publishing.

Sheelapriya, G. and Murugesan, R. (2017). Stock price trend prediction using Bayesian regularised radial basis function network model. *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, *46*(2), 189-211.

Shiller, R. J. and Pound, J. (1989). Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization*, *12*(1), 47-66.

Shmueli, G. and Koppius, O. R. (2011). Predictive analytics in information systems research. *MIS Quarterly*, 35(3), 553-572.

Siegel, J. (2006). The 'noisy market' hypothesis. *Wall Street Journal*, *14*, A14.

Siekmann, S., Kruse, R., Gebhardt, J., Van Overbeek, F. and Cooke, R. (2001). Information fusion in the context of stock index prediction. *International Journal of Intelligent Systems*, *16*(11), 1285-1298.

Simon, H. A. (1990). Prediction and prescription in systems modeling. *Operations Research*, *38*(1), 7-14.

Singh, J. and Khushi, M. (2021). Feature Learning for Stock Price Prediction Shows a Significant Role of Analyst Rating. *Applied System Innovation*, *4*(1), 17.

Singh, R. and Srivastava, S. (2017). Stock prediction using deep learning. *Multimedia Tools and Applications*, *76*(18), 18569-18584.

Spann, M. and Skiera, B. (2003). Internet-based virtual stock markets for business forecasting. *Management Science*, *49*(10), 1310-1326.

Spooner, M. C. (1984). Origin of fundamental analysis. *Financial Analysts Journal*, *40*(4), 79-80.

Srivastava, S. (2017). Tradeoff and Forecasting Approach for Social Responsible Investing: Artificial Neural Network. *Vision*, *21*(2), 143-151.

Steyn, Dimitri H. W.; Greyling, Talita; Rossouw, Stephanie; Mwamba, John M. (2020) :Sentiment, emotions and stock market predictability in developed and emerging markets, GLO Discussion Paper, No. 502, Global Labor Organization (GLO), Essen.

Sudhakaran, S. and Balasubramanian, P. (2016, September). A study on the impact of macroeconomic factors on S&P BSE Bankex returns. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2614-2618). IEEE.

Sujatha, K. V. and Sundaram, S. M. (2010a). A combined PCA-MLP model for predicting stock index. In *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India* (pp. 1-6).

Sujatha, K. V. and Sundaram, S. M. (2010b). Stock index prediction using regression and neural network models under non normal conditions. In *INTERACT-2010* (pp. 59-63). IEEE.

Tan, T. Z., Quek, C. and Ng, G. S. (2007). BIOLOGICAL BRAIN-INSPIRED GENETIC COMPLEMENTARY LEARNING FOR STOCK MARKET AND BANK FAILURE PREDICTION[1]. *Computational Intelligence*, *23*(2), 236-261.

Thakkar, A. and Chaudhari, K. (2021a). Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions. *Information Fusion*, *65*, 95-107.

Thakkar, A. and Chaudhari, K. (2021b). A Comprehensive Survey on Deep Neural Networks for Stock Market: The Need, Challenges, and Future Directions. *Expert Systems with Applications*, 114800.

Thenmozhi, M. (2006). Forecasting stock index returns using neural networks. *Delhi Business Review*, *7*(2), 59-69.

Tirunillai, S. and Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science*, *31*(2), 198-215.

Treleaven, P., Brown, R. G. and Yang, D. (2017). Blockchain technology in finance. *Computer*, *50*(9), 14-17.

Vardhan, H., Sinha, P. and Vij, M. (2015). Behavior of Indian sectoral stock price indices in the post subprime crisis period. *Journal of Advances in Management Research*, *12*(1), 15-29.

Venkatesh, V., Bala, H. and Sambamurthy, V. (2016). Implementation of an information and communication technology in a developing country: A multimethod longitudinal study in a bank in India. *Information Systems Research*, *27*(3), 558-579.

Verma, I., Dey, L. and Meisheri, H. (2017, August). Detecting, quantifying and accessing impact of news events on Indian stock indices. In *Proceedings of the International Conference on Web Intelligence* (pp. 550-557). ACM Digital library.

Von Szeliski, V. (1956). Predicting Stock Market Trends by Structure Analysis. *Financial Analysts Journal*, *12*(2), 51-61.

Walczak, S. (2001). An empirical analysis of data requirements for financial forecasting with neural networks. *Journal of Management Information Systems*, *17*(4), 203-222.

Wang, H., Lu, S. and Zhao, J. (2019). Aggregating multiple types of complex data in stock market prediction: A model-independent framework. *Knowledge-Based Systems*, *164*, 193-204.

Wang, J. J., Wang, J. Z., Zhang, Z. G. and Guo, S. P. (2012). Stock index forecasting based on a hybrid model. *Omega*, *40*(6), 758-766.

Weng, B., Ahmed, M. A. and Megahed, F. M. (2017). Stock market one-day ahead movement prediction using disparate data sources. *Expert Systems with Applications*, *79*, 153-163.

Weng, B., Martinez, W., Tsai, Y. T., Li, C., Lu, L., Barth, J. R. and Megahed, F. M. (2018). Macroeconomic indicators alone can predict the monthly closing price of major US indices: Insights from artificial intelligence, time-series analysis and hybrid models. *Applied Soft Computing*, *71*, 685-697.

White, H. (1988). Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Return. In *Proceedings of International Conference on Neural Networks.* IEEE (pp. II-451).

Wold, H. (1954). Causality and econometrics. *Econometrica: Journal of the Econometric Society*, 162-177.

Wu, J. L. and Hu, Y. H. (2012). Price–dividend ratios and stock price predictability. *Journal of Forecasting*, *31*(5), 423-442.

Xie, M., Li, H. and Zhao, Y. (2020). Blockchain financial investment based on deep learning network algorithm. *Journal of Computational and Applied Mathematics*, 112723.

Xu, S. X. and Zhang, X. (2013). Impact of Wikipedia on market information environment: evidence on management disclosure and investor reaction. *MIS Quarterly*, *37*(4), 1043-1068.

Zeinalizadeh, N., Shojaie, A. A. and Shariatmadari, M. (2015). Modeling and analysis of bank customer satisfaction using neural networks approach. *The International Journal of Bank Marketing*, *33*(6), 717.

Zhao, J. L., Fan, S. and Yan, J. (2016). Overview of business innovations and research opportunities in blockchain and introduction to the special issue. *Financial Innovation*, *2*(1), 1-7.

Ziegeldorf, J. H., Matzutt, R., Henze, M., Grossmann, F. and Wehrle, K. (2018). Secure and anonymous decentralized Bitcoin mixing. *Future Generation Computer Systems*, *80*, 448-466.

Zorn, T., Dudney, D. and Jirasakuldech, B. (2009). P/E changes: some new results. *Journal of Forecasting*, *28*(4), 358-370.

PUBLICATIONS IN PH.D. PROGRAMME

International Journals

1. Arjun R, Abhisek Kuanr, Suprabha KR (2021), Developing banking intelligence in emerging markets: Systematic review and agenda, *International Journal of Information Management Data Insights*, Volume 1, Issue 2, 100026, ISSN: 2667-0968 (DOAJ). DOI: https://doi.org/10.1016/j.jjimei.2021.100026 URL: https://www.sciencedirect.com/science/article/pii/S2667096821000197

2. Arjun, R.S. and Suprabha, K.R. (2020). A Bibliometric Review of Stock Market Prediction: Perspective of Emerging Markets. *Applied Computer Systems*, *25*(2), 77-86. A Journal of Riga Technical University. eISSN: 2255-8691. DOI: https://dl.acm.org/doi/abs/10.2478/acss-2020-0010 [Emerging Source Citation Index (ESCI), JCI= 0.16, ACM Digital Library, JUFO Finland ranking level 1].

3. Arjun, R. and Suprabha, K. R. (2020). Innovation and Challenges of Blockchain in Banking: A Scientometric View. *International Journal of Interactive Multimedia & Artificial Intelligence*, *6*(3). pp. 7-14. DOI: 10.9781/ijimai.2020.03.004, ISSN: 1989-1660 [Science Citation Index-Expanded with IF = 3.137, Q2 in Clarivate JCR June 2021, Scopus]

4. Arjun, R. and Suprabha, K. R. (2019). Forecasting banking sectors in Indian stock markets using machine intelligence. *International Journal of Hybrid Intelligent Systems*, *15*(3), 129-142. ISSN: 1448-5869, eISSN: 1875-8819. [EBSCO, DBLP, ACM Digital Library indexed, CORE journal rank C]. DOI: 10.3233/HIS-190266

Book Chapters

1. Arjun, R., Nishmitha N. and Suprabha K. R.. (2021). Financial Technology Implications: Emerging Markets Context. In Management Association, I. (Ed.), *Research Anthology on Concepts, Applications, and Challenges of FinTech* (pp. 239-260). IGI Global. 10.4018/978-1-7998-8546-7.ch014 [Publisher reprint]

2. Arjun, R., Suprabha K.R. and Majhi R. (2021). Deep Learning for Stock Index Tracking: Bank Sector Case. In: Bhateja V., Peng SL., Satapathy S.C., Zhang YD. (eds) *Evolution in Computational Intelligence*. Advances in Intelligent Systems and Computing, vol 1176. Springer, Singapore. DOI: 10.1007/978-981-15-5788-0_29 [Web of Science, Scopus]. Print ISBN: 978-981-15-5787-3.

3. Arjun, R. and Suprabha K.R. (2020). Modeling Hybrid Indicators for Stock Index Prediction. In: Abraham A., Cherukuri A., Melin P., Gandhi N. (eds) Intelligent Systems Design and Applications. ISDA 2018. *Advances in Intelligent Systems and Computing*, vol 940. Springer, Cham. DOI: 10.1007/978-3-030-16657-1_18 (WoS, Scopus, CORE- C). ISSN: 2194-5357.

4. Arjun, R., Nishmitha N. and Suprabha K. R., (2020). Financial Technology Implications: Emerging Markets Context. In Anshari, M., Almunawar, M. N., & Masri, M. (Ed.), *Financial Technology and Disruptive Innovation in ASEAN* (pp. 34-62). IGI Global. ISBN: 9781522591832. DOI: 10.4018/978-1-5225-9183-2 [Invited article]

Working Preprint

Arjun, R. and Suprabha, K.R. 2018. “Predictive modeling of stock indices closing from web search trends" Papers 1804.01676, *arXiv.org*. (Cornell University, RePEC indexed) https://ideas.repec.org/p/arx/papers/1804.01676.html [1 SCI citation]

Manuscripts under review/communicated

1. Extended abstract titled “Stock Prediction Framework for Expert Investment Recommendation system”. Accepted for oral presentation for 2020 International Conference on Partial Least Squares Structural Equation Modeling in Beijing, China (PLS 2020). Postponed to October 28-30, 2022. Springer (Scopus).

Referred Conference Proceedings

1. Arjun, R & D'Souza, S.C. (2016). Software Analytics Platform for Converged Healthcare Technologies. *Procedia Technology*, *24*, 1431-1435. Presented at the ICETEST 2015 held at GEC Thrissur, Kerala. DOI: http://dx.doi.org/10.1016/j.protcy.2016.05.169, ISSN: 2212-0173. Accession WOS: 000387696400187 (Elsevier, Web of Science). https://www.sciencedirect.com/science/article/pii/S2212017316302596

2. Arjun, R. & Sunil, C. D'Souza (2015). Mobile Health for Inclusive Growth among Rural Indian Population in *International Conference on Inclusive Growth and Profits with Purpose: New Management Paradigm* held at IIM Bangalore organized by IIM-B & International Management Research Academy, U.K. ISBN: 978-0-9573841-3-2.(Competitive Paper, abstract only)

Blogs

- "*Impact making research or scientific trash? Crossroads ahead*" in the Springer Nature Behavioural and Social Sciences Journal. https://socialsciences.nature.com/posts/54674-impact-making-research-or-scientific-trash-crossroads-ahead

Workshops presented

1. Attended and presented at the doctoral symposium in the Author Development International workshop organized by the Association for Information Systems (AIS) India Chapter and University of Passau at IIT Madras, December 2019.
2. Research Talk for Ph.D. Student Colloquium titled "*Stock Prediction models: Indian context*" on International Workshop on Machine Intelligence on Data Science (MIDS 2019) organized by MIR Labs, USA on 31st May-June 1, 2019
3. Attended and presented research seminar at Workshop on Time-series Analysis using Artificial Neural Networks (TSANN), NIT Calicut. Dec 2018.
4. Brief research talk and workshop at the Advanced Research Methods in Healthcare Management Services at IIM Ahmedabad held on Sept 3-6, 2016.

Citations received

1. Liang, X., Ge, Z., Sun, L., He, M., & Chen, H. (2019). LSTM with wavelet transform based data preprocessing for stock price prediction. *Mathematical Problems in Engineering*. [SCI, Q3]
2. González, J. C., García-Díaz, V., Núñez-Valdez, E. R., Gómez, A. G., & Crespo, R. G. (2020). Replacing email protocols with blockchain-based smart contracts. *Cluster Computing*, *23*, 1795-1801. [SCI, Q1].
3. Casallas, J.A.T., Lovelle, J. M. C., & Molano, J. I. R. (2020). Smart Contracts with Blockchain in the Public Sector. *International Journal of Interactive Multimedia and Artificial Intelligence*, *6*(3). pp.63-72. [SCI, Q2]
4. Yang, J., Ma, X., Crespo, R. G., & Martínez, O. S. (2020). Blockchain for supply chain performance and logistics management. *Applied Stochastic Models in Business and Industry*. (SSCI. IF= 1.17, Q2, ABDC B ranked)
5. Aponte, F., Pineda, M., Gutierrez, L., Meriño, I., Salazar, A., & Wightman, P. (2020). Cluster-Based Classification of Blockchain Consensus Algorithms. *IEEE Latin America Transactions*, *19*(4). [SCI, Q4 journal]
6. Rahman, A., Islam, M. J., Rahman, Z., Reza, M. M., Anwar, A., Mahmud, M. P. & Noor, R. M. (2020). Distb-condo: Distributed blockchain-based iot-sdn model for smart condominium. *IEEE Access*, *8*, 209594-209609. [SCI, Q1]
7. Ishwarappa, Anuradha, J. (2021) Big data based stock trend prediction using deep CNN with reinforcement-LSTM model. *International Journal of System Assurance Engineering and Management*, 1-11 [Scopus]
8. Sun, X., Yang, T., & Hu, B. (2021). LSTM-TC: Bitcoin coin mixing detection method with a high recall. *Applied Intelligence*, 1-14. [SCI, Q2]
9. Sirohi, A. (2020) Relevance of Blockchain technology in the scenario of escalating cybercrime in banking sector in the UK. *Research Proposal*. Bournemouth University.
10. Kołodziej, M. (2021). Development Factors of Blockchain Technology Within Banking Sector. In *Contemporary Trends and Challenges in Finance* (pp. 125-138). Springer, Cham.
11. Martín, C., Quintana, D., & Isasi, P. (2021). Dynamic Generation of Investment Recommendations Using Grammatical Evolution. *International Journal of Interactive Multimedia & Artificial Intelligence*, *6*(6). [SCI, Q2]
12. Chiu, C. H., & Tsai, Y. C. (2021, March). Predicting Period Stock Spread Ranking Using Revenue Indicators and Machine Learning Techniques. In *IOP Conference Series: Earth and Environmental Science* (Vol. 704, No. 1, p. 012014). IOP Publishing. [Scopus]

BRIEF BIO-DATA

ARJUN R

Residence: S/o Shri. Somanathan. E
"Vaasudevam" (H), Seminary Road,
M.G.Kavu P.O, Thrissur, Kerala – 680 581
E-Mail: arjrs123[at]gmail[dot]com
Mobile: (+91) 9497656152

ORCID: http://orcid.org/0000-0002-2770-6164
ResearcherID: P-4860-2017
Publons: https://publons.com/researcher/1548325/arjun-r/
ResearchGate: https://www.researchgate.net/profile/Arjun_R

QUALIFICATIONS

- M.Tech specialized in Software Engineering from Cochin University of Science and Technology in June 2011, First Class and CGPA: 7.29/10.

  Master's Dissertation: Software Development for Launch Vehicle Telemetry Data Management (Project in VSSC, ISRO)
- B.Tech in Information Technology from Sun College of Engineering & Technology under Anna University, Chennai in June 2007. First Class and Aggregate: 65%. Term project: Intranet Mailing System with SSL.
- Diploma in Electronics and Communication Engineering from Sun Institute of Technology under Directorate of Technical Education, Tamil Nadu on May 2004, First class and Aggregate: 79.24%
- Secondary School Certificate from Arya Central School, Trivandrum, 2001 under Central Board of Secondary Education, New Delhi

SCHOLARLY ARTICLES

- Integrated Environment for Launch Vehicle Telemetry Data Management. In 2012 *International Conference on Data Science and Engineering* (ICDSE), (pp. 171-174). IEEE. INSPEC Accession Number: 12964077. ISBN: 978-1-4673-2146-4. IEEEXplore (Scopus). DOI: 10.1109/ICDSE.2012.6281894

WORKSHOPS PARTICIPATED

1. AICTE workshop on Blockchain, NITK, 2019
2. Predictive Analytics and its Applications, workshop at NITK, 2019
3. Applied Data Science & Business Analytics using R, at NITC, 2017
4. Algorithmic Foundations of Wireless Sensor Networks with Applications
5. Software Mining and Analysis, Advanced GIAN course at NITK.

SELECT CERTIFICATIONS

- Certificate of Excellence in Research data management course from Elsevier
- Web of Science Basic Series certification from Clarivate Analytics, 2019
- Statement of Accomplishment with Distinction in Statistical Learning (online course), Stanford University, 2018
- Certificate of Achievement in course: Big Data Analytics, Griffith University
- Certificate of Achievement in the course: Being a researcher in Information Science and Technology, Politecnico di Milano, Italy
- Early Career Learning Program of Research Impact, Taylor and Francis
- MATLAB for Financial Applications, MathWorks, 2019
- Focus on Peer Review, Master class online course by Springer Nature, 2019
- Certificate of completion in Google Analytics from Google LLC.
- IEEE certification on Mobile Web Technologies for the Developing World

PROFESSIONAL ASSOCIATIONS

- ACM membership. 8259524
- IEEE membership no. 91142019
- Academy of Management member no. 140134

SERVICES

- Technical program committee member and reviewer in 19th ICCSA 2019 at St. Petersburg University, Russia and reviewer for FICTA 2020, ICIS 2021.
- Reviewed for paper submissions to Academy of Management Annual Meeting on August 7-11, 2020, Vancouver, Canada & AoM 2021 virtual annual meeting.
- Editorial Board member for International Journal of E-Business Research (IJEBR), IGI Global Publishers, USA (ESCI, Scopus and Chartered ABS l).
- Reviewer for EAI Transactions on Scalable Information Systems, Information Resources Management journal, Cogent Business and Management journal.

End of Thesis