Predicting stock returns of banking sector using machine learning models: Evidence from South Asian Economies

[1]Ghulam Nabi, [2]Javid Iqbal

[1]PhD student, COMSATS University Islamabad., fa19-pms-003@isbstudent.comsats.edu.pk
[2]Assistant Professor, COMSATS University Islamabad., javidiqbal@comsats.edu.pk

ABSTRACT

Banks are very important financial intermediaries in the financial system of a country. The objective of the study is to predict the stock market returns of the banking sector in South Asia. Banks stock return movements may have significant practical ramifications for a variety of decision makers, including current shareholders, potential investors, peer bank managers, and credit rating agencies. This study integrates textual data with financial information from leading news articles to predict stock returns in the banking sector of South Asian Economies with daily observations of 231,458 covering period from January 05, 2010 to January 12, 2024 by applying neural network technique. The study's findings showed that the news articles are important predictor of stock returns of South Asian banking sector. The findings also revealed that the information in news articles in news papers provide important information that may be used for stock returns. Therefore, investors, policymakers and researchers can leverage textual information in business news articles with financial information for better and well-informed decision-making. Investors can concentrate on potential investment opportunities, avoiding market risks and establishing portfolios in South Asia as bank stock is generally considered as an index of future economic expansion.

Keywords: *News Articles, Stock Market Returns, Banking Sector, South Asian Economies, Machine Learning*

INTRODUCTION

The two most important components of the financial system are banks and the stock market which are also linked together. Important capital market investment operations are connected to the banking industry (Chen et al., 2022). The capital markets and economic activity directly indicate the banking system's health (Armagan, 2023). The banking sector has a close nexus with the other sectors of the economy. Banks are distinct from non-financial companies because of their complex capital structure, business operations, governance structure, and regulations (de Haan & Vlahu, 2016). Bank failures impact other financial systems, putting the entire economy at risk (Levine,2012). Macroeconomic externalities are more likely to occur when big banks fail (Boyd & Runkle,1993). Banks' crucial role is providing credit, which helps ensure all sectors have the money they need.

Likewise, shares of banks listed on the stock market are traded similarly to other industries, offering trading opportunities to investors and funding sources for the banks. According to Chen et al., 2022), the banking sector has a strong hold on the stock market, which forms the basis of the capital market. Banking is involved in important capital market investing activities (Chen et al., 2022). According to Groenewold et al. (2003), market efficiency is increased by trading bank shares. An empirical study of the Chinese stock market found that the exclusion of banks hinders the stock market's efficiency. Mittal and Garg (2021), converged that bank stocks play a significant role in GDP growth, allowing investors to concentrate on investment opportunities, particularly in emerging markets. Sensarma and Jayadev (2009) described that bank stock returns are correlated with banks' ability to manage risk. Improved risk management ratings or competencies help banks' stock market returns and can benefit investors. Hence, stock returns volatility of banking securities is one of the most widely utilized metrics for evaluating bank risks (Dogra et al., 2021). Due to its significance, predicting bank stock prices is a crucial topic in the investment area (Chen et al., 2022). Limited evidence was found from previous research, which specifically targets the stock price prediction of the banking sector (Arjun & Suprabha, 2019). Malik et al. (2017) explain that bank stocks have historically been considered profitable equity investments in stock exchanges/capital markets. According to Wagle (2021), beyond supply and demand, additional factors also impact commercial bank stock prices, which makes predicting the price of the banking industry's stock crucial for investment strategies.

The stock market is affected by non-financial information such as news information in newspaper articles, online news sources, and other social media (Seng & Yang, 2017; X. Yao et al., 2023; Yuan et al., 2020). The stock market return prediction is challenging as many factors influence the stock market. The news articles contain information which are related to changes in economic indicators, political turmoil, various global events, policies concerning investment strategies and financial conditions announced by the organization or government (Lin et al., 2022).

According to Nemes & Kiss (2021), researchers and investors have already started to examine the relationship between news headlines and firms' stock prices. Wojarnik (2021) demonstrated an association between stock price behavior and information on Twitter or other social media platforms. Some other studies also suggested that textual information in the form of financial news can help accurately predict stock prices (Dogra et al., 2021a; Lin et al., 2022; Nemes & Kiss, 2021a; Wojarnik, 2021; Yuan et al., 2020).

There is a body of literature that studies stock market prediction. This study focuses on predicting stock returns of the banking sector with the help of novel approach by using sentiment in news articles typically reported in daily newspaper business sections as evidenced from the literature. Therefore, a major research gap that has not been addressed in previous studies is the use of textual information from news articles and financial data by applying advanced machine learning models to predict the stock returns of the banking industry. These gaps have not been addressed in the stock market research literature.

Our results suggest that the news articles are important predictor of stock returns of South Asian banking sector. The findings also revealed that the information in news articles in newspapers provide valuable information that may be used for predictions of stock returns.

This study contributes to existing literature in the following ways. Firstly, the study contributes by determining the relationship between news articles and changes in the banking industry's stock price. Secondly, we focus on large qualitative datasets, i.e., news articles to extend the stock returns prediction of the banking sector. Hence, investors and regulators can take advantage of the illuminating capabilities of integrated textual information along with financial data, which in turn can improve their decision-making. Finally, the study results provide a valuable contribution by analyzing the banking sector of South Asian countries.

The remainder of the paper is organized in this manner. Section 2 presents the review of literature. Section 3 provides theoretical background and hypotheses. Section 4 provides research methodology, and Section 5 presents the results. The final section concludes the discussion.

LITERATURE REVIEW

Banks, stock exchanges, and other depository financial intermediaries are essential elements that make up the financial system (Thakor, 1996). Literature has witnessed that there is a strong connection between financial sector and economic growth (Beck et al., 2000; King & Levine, 1993; Levine & Zervos, 1998; Odhiambo & Nyasha,2022; Zou et al., 2022). There is an abundance of literature arguing that capital accumulation and the level of financial intermediary development are the key factors for economic growth, such as (King & Levine, 1993; Levine, 1991; Levine & Zervos, 1998; Peia & Roszbach, 2015; Shen, 2013; Neusser & Kugler,1998; Rousseau and Wachtel,1998; Demirguc- Kunt and Maksimovic,1998). King and Levine (1993) employed data from 80 countries between 1980 and 1989 to examine the relationship between financial development and economic expansion. The research findings indicated a robust association between financial development and economic expansion. A robust relationship was discovered between the expansion brought about by the actual building up of capital and the effective distribution of resources. Asian Development Bank, (2009) reported that most South Asian countries' financial systems are currently dominated by banks. Empirical evidence confirmed the strong relationship between the banking sector and the stock market. As a crucial component of the financial system, stock markets aggregate the funds from several smaller savers and allocate them to the most critical uses (Stiglitz et al., 1998). Samarasinghe and Uylangco (2021) examined the effects of total stock market liquidity on the conventional source of bank business between 1999 and 2014 using data from 7279 institutions in 39 countries. The study's findings demonstrated that a general increase in stock market liquidity substantially negatively impacts bank deposits and lending rates. The relationship between the liquidity of the stock market and the origin of bank deposits and loans is contingent upon the growth rate of the country and the degree of investor protection afforded by the market. Biswaset al., (2017) Using monthly data from 2006 to 2015, canonical correlation analysis was used to investigate the structural dependency between Bangladesh's stock market and banking sector trends. It was concluded that the two financial systems have developed independently throughout Bangladesh's financing of economic activities, as evidenced by the positive but not statistically significant correlation between the banking sector and the development of the stock market. Samarasinghe, (2023) conducted an empirical analysis to examine the relationship between bank stability and the stock market. According to the study, banks tended to diversify their business operations by depending on non-traditional sources of income as stock market liquidity increased. Banks used the stock market to compensate for losses from performing their regular intermediary duties. The study concluded that, especially in developed nations, the stability of banks and stock market liquidity are positively correlated.

There is a rich body of literature available for predicting the performance of the stock market; however, no literature specifically addressed the stock performance of the banking sector of South Asian Countries. Most past research on predicting stock performance only covered financial variables without addressing textual data derived from the newspaper articles. Therefore, authors of the most recent studies proposed that textual data and investor sentiment should be utilized in addition to financial data to predict stock market return. Further there was scant literature using sentiment analysis for stock market return prediction of banking sector of south Asian countries particularly capturing the sentiments of the investors while making decision for investment allocation. Newspaper articles often contain important textual information about companies and industries that have severe impact on a company's or sector's stock price (Dogra et al., 2021). According to Lin et al. (2022), few studies evaluated the movement in stock prices by considering text data and stock price data, which makes it difficult to predict stock values due to a variety of factors, including nonlinearity, noisy data, political turmoil, economic fluctuations, and global events usually covered in text form, such as news articles. Therefore, analyzing textual data to forecast stock market activity was crucial.Li and Pan, (2022) conducted a study on the prediction of stock prices based on financial and textual news data. Stock prices are affected due to different factors, which may be company-related and industry-specific, such as declarations of dividends, the arrival of an innovative product or a product recall, the acquisition of an important deal, reductions in staff, a big management change, an anticipated takeover or merger, and financial scandals or blunders. The study results showed that some news stories have a prolonged impact on market values. Nyakurukwa and Seetharam (2023) analyzed the influence of sentiment analysis on price variations of dual-listed stocks to assess how news on Twitter and stock prices are connected. The Anglo-American and Anglo-Australian twin stocks exhibit a negative link, whereas the Anglo-South African twin stocks are positively associated with the news sentiment.

Seng and Yang, (2017) examined the connection between financial news and stock price volatility. The experiment's findings showed a statistically significant correlation between stock market volatility and financial news. Additionally, it has been discovered that positive stock returns have an adverse correlation with good news and a positive correlation with the news's supplementary information score. Zhang et al. (2022) investigated using a multi-module feature fusion algorithm based on pre-trained language to forecast long-term stock movement. This method included textual features from Chinese research reports instead of investor sentiments, which were primarily focused on recent studies. The proposed model has the greatest performance of 79.2% in the one-year stock movement prediction task. The findings showed that research reports with basic data were more effective and were crucial for long-term stock forecasts.

METHODOLOGY

Sample

Initially, we select all the banks of South Asian Countries. Asian Development Bank (2009) reported that banking sector is mainly dominated by South Asian countries' financial systems. Further, we select schedule banks of four South Asian countries, which share similarities i.e. Bangladesh, India, Pakistan and Sri Lanka. Financial data is retrieved from Yahoo Finance and DataStream, which includes historical bank stock prices and other financial variables. For construction of sentiment index, news articles are scraped from the website of business sections of the leading national newspapers using Python programing. Our final sample consists of 235,864 daily observations from 122 banks in four countries covering the period January 05, 2010 to January 12, 2024. The number of banks and time duration varies due to the availability of news articles and financial variables.

Variables and their details

Outcome variable

Determining the effect of news articles on the stock return of the banking industry in South Asian economies is the main goal of the study. Thus, the outcome variables in this study is stock returns of banks. The literature has extensively used this variable (Aayale et al., 2022; Arjun & Suprabha, 2019; Y. Chen et al., 2022a; Q. Liu et al., 2022; Sharma et al., 2022). We use following formula to calculate the stock returns of banking sector in equation 1:

$$Return = \frac{ClosePrice - OpenPrice}{OpenPrice} * 100 \qquad (1)$$

Input Variables

In addition to the dependent variable, we employ 7 financial variables, including the sentiment index serving as the primary independent variable.

Sentiment Index of Textual News Data

The sentiment index developed from news articles is the primary independent variable to predict stock returns of banking sector in South Asian Economies. This index has been frequently used in the literature (Feldman et al., 2008; C. Liang et al., 2020). To extract positive and negative words from news articles, we employ Harvard Sociopsychological dictionary to develop sentiment index.

We used sentiment index formula developed by (Uang et al., 2006) and has been widely used by many other researchers (Henry, 2008; Iqbal & Riaz, 2022). This is calculated in equation 2.

$$\text{Sentiment Index} = \frac{Positive\ Words - Neagtive\ Words}{Positive\ Words + Neagtive\ Words} \qquad (2)$$

The sentiment index is +1 for completely positive and -1 for completely negative.

Control Variables

Apart from our main variable i.e. sentiment index developed from news articles using Harvard Dictionary, we use other control variables. These variables include earnings per share (EPS), Return on Equity (ROE), Return on Assets (ROA), Price to Book Value Ratio (P/B Ratio), Price to Book Value Ratio (P/B Ratio), Dividend Yield, and Total Asset. These variables have been widely used in the previous research to link with stock returns prediction (Aayale et al., 2022; Al Nimer et al., 2015; Arjun & Suprabha, 2019; Liu et al., 2022; Sari, 2021; Sharma et al., 2022; Sidhu et al., 2022).

Artificial Neural networks (ANN)

Artificial Neural networks (ANN) or Neural Networks (NN) is a biologically inspired machine learning technique, which is used for predictions. This technique yields higher prediction accuracy when compared to traditional econometric methods (Sezer et al., 2020). The underlying algorithm is first trained using a labeled training dataset and then unseen test data is provided for prediction.

Neural network's Architecture

An input layer, an output layer, and neurons in a hidden layer make up an artificial neural networks (ANN) architecture. Neuron is the basic processing unit of a neural network. According to Anderson and Mcneill (1992), a biological neuron receives information from multiple sources, integrates it in some fashion, applies a nonlinear operation to the result, and then outputs the result. The layer containing all feature variables denoted by x1, x2, x3, and so on up to x_n is known as the input layer (Chhajer et al., 2022). The general neural network's structure is given in figure 1.

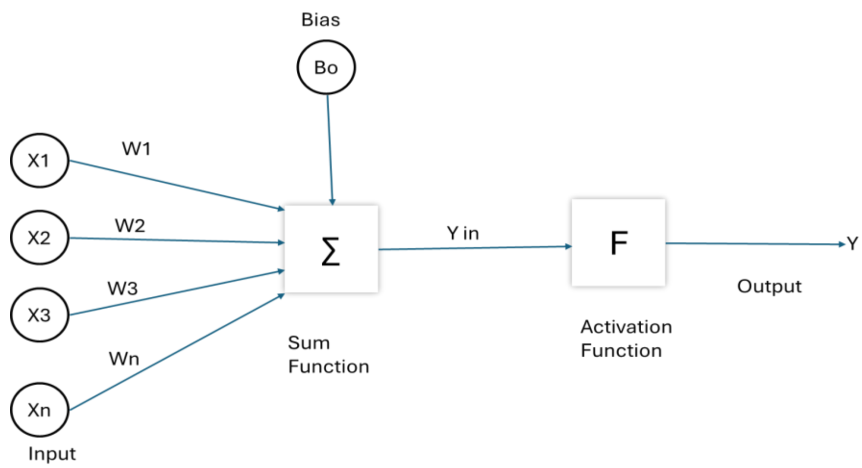

Figure 1: Architecture of neural networks.

The functions and input weights are described by the following equation (3).

$$0(SR) = f\left(b_0 + \sum_{i=1}^{n} W_i Z_{i,t-1}\right) = f(b_0 + W_i Z_{i,t-1}) \qquad (3)$$

$$(\, i = 1, \ldots\ldots\ldots, \; N, t = 1, \ldots\ldots\ldots, T)$$

Where O(SR) represents the output variable as stock return calculated by a neural network, b_0 denotes the bias of the output neuron, $W = w_{1,} \ldots \ldots .. w_n$ are the weights of input variables and Z is the vector of values of all the input variables of all the variables including sentiment index and other financial variables.

The complexity of the function depends on how many hidden layers or hidden neurons are added or eliminated. The dataset is split into the test dataset and the training dataset for predictive analysis. The algorithm is trained on the training dataset, after which it looks for patterns in the data and modifies the weights assigned to each input variable. A network model is fitted to the test dataset for generalization of results.

RESULTS AND DISCUSSION

Descriptive Statistics

We used Harvard Sociopsychological dictionary to develop sentiment index. The mean value of sentiment index is 0.27 ranging from -1 to 1 which also suggests that financial news of articles have impact on stock value of the banks in South Asian region and that investor opinions are favorable.

Despite good financial fundamentals, negative news published in articles affected market perception and added to the negative stock return. The sentiment variable's mean value suggests that investors may rely on financial factors along with other information sources considering when making decisions about their investments rather than consulting other information sources. Moreover. individual sentiments negatively influenced to bank stock returns instead of market sentiment. Descriptive statistics also reveal that the mean value of return on equity is high compared to the mean return on assets, which implies that the banking sector in South Asia is highly leveraged. The sector's poor performance is indicated by the banks' inability to utilize their assets profitably, as evidenced by the low mean value of return on assets despite the high mean value of total assets. Banks usually have huge assets as they hold substantial quantities of loans, securities, and other financial instruments, ROA calculates a bank's efficiency to convert a profit from its assets. Even with a high total asset base, a low return on assets (ROA) suggests that the bank might not be making sufficient funds for its size.

Table 1: statistics of sentiment index and financial variables

Variable	Mean	Median	Std Dev	Min	Max	25th Pctl	75th Pctl	95th Pctl
Market Return	-0.04	0	3.95	-97.2	1030.39	-1.18	0.87	4.02
Sentiment_Harvard	0.27	0.26	0.14	-1	1	0.18	0.36	0.5
EPS	5.22	1.87	13.31	-84.58	96.01	0	7.61	26.89
ROA	0.96	0.96	0.86	-5.03	4.9	0.7	1.27	2.29
ROE	8.79	8.79	11.66	-144.79	53.27	7.69	13.87	20.73
Log_Assets	20.65	20.8	1.58	6.21	24.81	19.64	21.58	22.81
Dividend Yield	3.39	3.39	3.32	0	33.47	1.03	3.39	10
Price-to-Book	1.2	1.2	0.85	-2.23	18.88	0.81	1.2	2.53

The price to book ratio compares the stock market and book values. According to the results shown in Table 1, the average book price is 1.2, which is higher than 1. This indicates that the stock of the banking industry is overpriced and has historically fared well. Typically, investors stay away of overpriced businesses and industries. The P/B ratio and ROE are both trending in the same direction when we combine them, indicating that the company's stock is overpriced, and the banking industry is not investing its profits in new initiatives. Thus, the outcome demonstrates that private banks' liquidity situation improved because of fund leverage. Extreme liquidity, however, would not be beneficial to a bank's performance if there was insufficient loan creation. High-capital banks may also have significant overhead costs without compromising their profitability. This can imply that the bank's cost structures, investment plans, or lending procedures are not as efficient as they could be. During this time, the global financial crisis and COVID-19 may have an impact on the banking sector in this area. The country's degree of development and macroeconomic factors may have an impact on the region's banking performance.

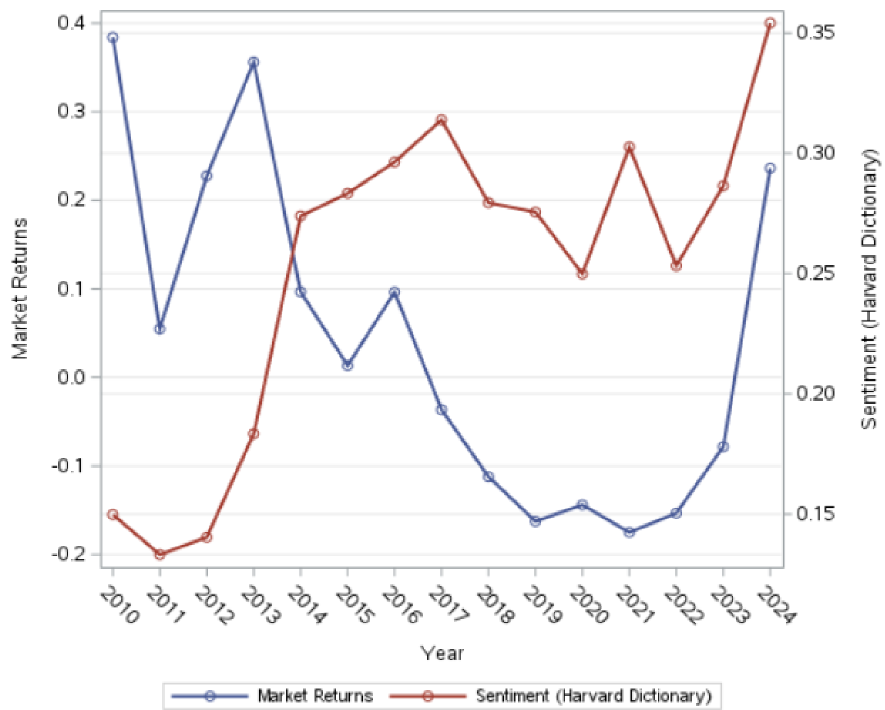

Figure 2: Market Returns and Sentiment Index of Harvard Dictionary

Figure 2 shows that market returns were declining in 2010 and abruptly increasing in 2012, is evidence of the similar trend in market returns observed from the sentiments generated by the Harvard Dictionary. This period also corresponds with the recovery of the global financial crises, which also has an impact on the South Asian region. In 2013, the sentiment line crosses over to the stock return line. The graph shows that the declining stock return trends persisted until 2021, at which point the market began to rebound until 2024. The findings indicate that investors' actions are significantly influenced by the financial information included in newspaper news items, as evidenced from the results portray in figures 2.

The above descriptive statistics further complemented with the analysis of correlation matrix as given in Table 2, which displays the findings of Pearson’s coefficient correlation analysis using the dependent and independent variables.

Out of the seven (07) independent variables, the results show that only Return on Assets (ROA) and Earnings Per Share (EPS) have a correlation (r =.682) over the 0.60 threshold level which means both are positively correlated. Furthermore, the Return on Equity (ROE) and Return on Assets (ROA) have a negative correlation with total assets, confirming that bank's size are not producing any returns and South Asian banks are not investing their money in lucrative ventures. Although they are below the threshold level, earnings per share (EPS) and return on equity (ROE-r=0.488) and return on assets (ROA-r=0.549) have a positive association, meaning that the higher the return on equity and assets, the higher the EPS. Price to book has a negative correlation with return on equity as well, confirming that the share market price is viewed as overpriced, a sign of the poor performance of the South Asian banking sector.

Table 2-Correlation matrix analysis

	Market Return	Sentiment Harvard	EPS	ROA	ROE	Total Assets	Dividend Yield	Price-to-Book
Market Return	1							
Sentiment_Harvard	-0.005	1						
EPS	0.004	-0.093	1					
ROA	0.022	-0.117	0.549	1				
ROE	0.020	-0.003	0.488	0.682	1			
Total_Assets	-0.009	-0.052	0.229	-0.068	-0.088	1		
Dividend Yield	0.016	-0.097	0.012	0.239	0.246	-0.164	1	
Price-to-Book	0.010	0.033	0.305	0.239	-0.078	0.046	-0.186	1

There is no substantial correlation between investor sentiment and any of the other independent variables. Although the sentiment variable has a negative association with market return, this correlation is not statistically significant, which further supports the findings of descriptive statistics showing that investors' investment decisions were adversely affected by the information in news items. The market return has a negative relationship with the bank’s asset sizes, but it is not a strong predictor.

Neural Network Results

Developing a network model is an iterative process. (Maier & Dandy, 2000). We scale the data from 0 to 1 to meet the activation function's requirements i.e., the hyperbolic tangent function, in order to optimize the neural network. We divide dataset into training and test dataset with the proportion of 70% and 30%, respectively (Haddi et al., 2013). When training the network, it's beneficial to start with one hidden layer and as many hidden neurons as half the total number of input and output variables (Ayodele, 2010). In this study, there were seven input variables and one output variable, resulting in nearly seven neurons in the hidden layer. We compute the MSE,

Table 3 **Predictive accuracy of neural network models using Harvard dictionary**

	Model	Train (70%) (N)	Neurons	Test (30%) (N)	*MSE*	*RMSE*	*SSE*
					Single hidden layer		
Panel 1	1	162000	2	69458	0.000022	0.004668	1.51232
	2	162000	4	69458	0.000022	0.004658	1.50529
	3	162000	6	69458	0.000022	0.00465	1.50060
	4	162000	8	69458	0.000021	0.00465	1.50280
					Two hidden layers		
			Neurons		***MSE***	***RMSE***	***SSE***
Panel 2	5	162000	2-1	69458	0.000022	0.004660	1.50860
	6	162000	4-2	69458	0.000022	0.004656	1.50444
	7	162000	6-3	69458	0.000021	0.004649	1.50000
	8	162000	8-4	69458	0.000022	0.004650	1.50024
					Three hidden layers		
			Neurons		***MSE***	***RMSE***	***SSE***
Panel 3	9	162000	4-2-1	69458	0.0000218	0.004678	1.51835
	10	162000	6-3-2	69458	0.0000216	0.004653	1.50213
	11	162000	8-4-2	69458	0.0000217	0.004657	1.50465
	12	162000	10-5-3	69458	0.0000216	0.004654	1.50304

RMSE, and SSE to find the deviation between the actual and anticipated values. Accurate predictions are made when these statistics are lower.

The accuracy of the models for three to ten neurons in a single hidden layer is shown in Table 3 Panel 1. Because having too few or too many neurons might lead to prediction bias, the number of hidden neurons in the hidden layer determines the accuracy of the model. Initially, we use 4 neurons in the network model for predicting stock returns. This provides us with 0.004658 RMSE and 1.50529 SSE. Further, we decrease the number of neurons by half of previous network i.e. 2 neurons to observe the accuracy of model. However, the accuracy was decreased. Hence, we decided to increase the neurons to improve model accuracy.

Table 3's findings show that a drop in prediction accuracy is correlated with an increase of neurons. In order to fit the model, we therefore add more neurons to the hidden layer. The network results of two hidden layers with many neurons arranged in various combinations are given in Table 3 under Panel-2. Out of all the neuron combinations in the test dataset, Network 5, with its two hidden layers and two neurons in layer 1, had the highest prediction accuracy because the test dataset has an lowest MSE of 0.00022, RMSE of 0.0046, and an SSE of 1.5. This is also presented in Figure 3.

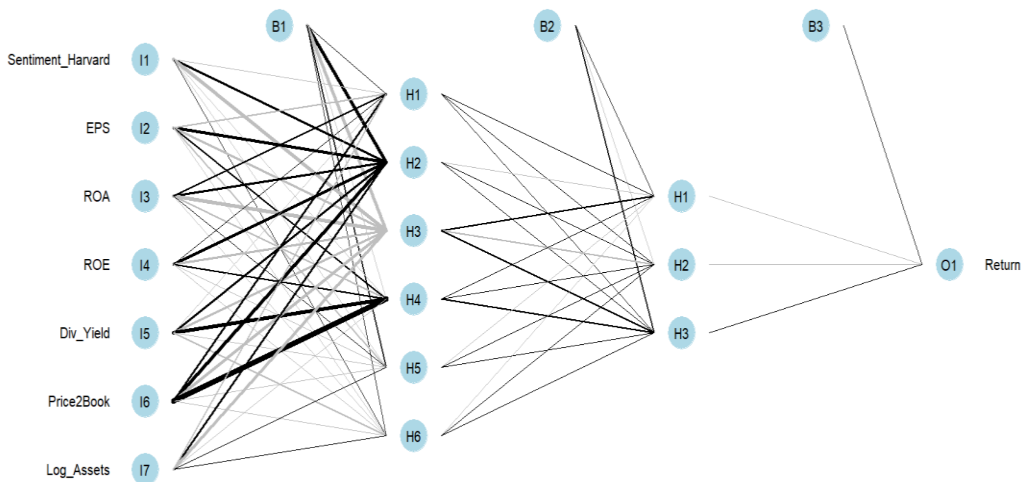

Figure 3: Neural Network with highest predictive accuracy

As the number of hidden layers rises in an attempt to raise the prediction rate, the model progressively approaches deep learning. There is an additional layer that has multiple neuron pairings. Table 3, Panel 3 displays the three hidden layers' predicted accuracy; It can be analyzed from Panel-3 that no significant change in network's error rate than that of all prior networks. As no change with more hidden layers, adding neurons and layers indicates the higher prediction accuracy with less error rate. The model, including three hidden layers, has the lowest prediction accuracy. Thus, we deduce that the network model with two hidden layers yielded more accurate predictions.

Significance of Variables in Neural Network Models

Since it can anticipate bank stock returns with the highest predictive accuracy, we select with neural network model 5. Sentiment indexes, such as +1 for positive sentiment and -1 for completely negative sentiments, are useful tools for gauging investor mood and appear to be crucial in predicting bank stock returns. The future stock price of banks may be positively or negatively correlated with this index.

Our results are consistent with previous studies on non-financial firms, which suggest that sentiments in news articles is an important predictor of stock returns (Nemes and Kiss, 2021; Dogra et al., 2021; Li and Pan, 2022; Seng and Yang, 2017; Yao et al., 2023). These studies provide that different news or events' effects on the price movements. Adding textual information to the model increases forecasting accuracy because news articles contain information related company-related and industry-specific, such as declarations of dividends, the arrival of an innovative product or a product recall, the acquisition of an important deal, reductions in staff, a big management change, an anticipated takeover or merger, and financial scandals. These news articles also contain macro-economic matters related to IMF, World Bank loans, pandemic, or related issues or global financial crisis.

CONCLUSION

The financial system consisted of financial markets (e.g., primary, secondary markets), and banks (e.g., commercial, specialized banks). Financial markets include primary markets that enable corporations to raise funds for their business ventures through public offerings, and secondary markets. The actual economy's operations rely on the stock market's efficient operation. Likewise, shares of banks listed on the stock market are traded similarly to other industries, offering trading opportunities to investors and funding sources for the banks.

This study integrates textual data with financial information from leading news articles to predict stock returns in the banking sector of South Asian Economies covering the daily observations January 05, 2010, to January 12, 2024. This study has used neural network model for better prediction accuracy.

Despite having large size by holding total assets, the sector's stock performance is poor, suggesting that South Asia's banking industry is heavily indebted. The banks' inability to make profitable use of their assets is a sign of the sector's bad performance. Regardless of the individual performance of the banks in the region, economic indicators and other global factors may have a detrimental impact on the region's economy.

To predict the financial performance of stock returns of banking sector, we integrate textual data with financial information from leading news articles to predict stock returns in the banking sector of South Asian Economies. For more accurate prediction, we use artificial neural networks. The study's findings showed that the news articles are important predictor of stock returns along with other financial variables of South Asian banking sector. The findings also revealed that the information in newspaper articles provides important information that may be used for stock returns. Therefore, investors, policymakers and researchers can leverage textual information in business news articles with financial information for better- and well-informed decision-making. Since bank stock is typically regarded as a predictor of future economic growth, the study's conclusions are concerning investors and policymakers who want to minimize market risks and build portfolios in South Asia. To identify the specific causes for the banking industry's poor performance with the economy's most important component, further study on the banking sector in South Asia might be conducted after including macroeconomic factors in addition to textual and financial components.

REFERENCES

Aayale, J., Seffar, M., & Koutene, J. (2022). View of Financial Indicators, Stock Prices and Returns_ Evidence from Banks Listed on the Stock Exchange of an Emerging Market. *International Journal of Accounting, Finance, Auditing, Management & Economics*, *3*(2–2), 533–551.

Arcuri, M. C., Gandolfi, G., & Russo, I. (2023). Does fake news impact stock returns? Evidence from US and EU stock markets. *Journal of Economics and Business*. https://doi.org/10.1016/j.jeconbus.2023.106130

Arjun, R., & Suprabha, K. R. (2019). Forecasting banking sectors in Indian stock markets using machine intelligence. *International Journal of Hybrid Intelligent Systems*, *15*(3), 129–142. https://doi.org/10.3233/his-190266

Asian Development Bank. (2009). *South Asia Economic Report Financial Sector in South Asia: Recent Developments and Challenges Printed in the Philippines*. www.adb.org

Asian Economic Integration Report 2023. (2023). https://doi.org/10.22617/TCS230031-2

Ayodele, T. O. (2010). *New Advances in Machine Learning*. IntechOpen. https://doi.org/https://doi.org/10.5772/225

Beck, T., Levine, R., & Loayza, N. (2000). Finance and sources of growth. *Journal of Financial Economics*, 261–300.

Biswas, S., Hossain, A., Podder, A. K., & Hossain, Md. N. (2017). A Canonical Analysis on the Relationship between Banking Sector and Stock Market Development in Bangladesh. *International Journal of Economics and Finance*, *10*(1), 167. https://doi.org/10.5539/ijef.v10n1p167

Brooke, J. (2009). *A SEMANTIC APPROACH TO AUTOMATED TEXT SENTIMENT ANALYSIS*. http://ir.lib.sfu.ca/handle/1892/112

Chen, Y., Wu, J., & Wu, Z. (2022). China's commercial bank stock price prediction using a novel K-means-LSTM hybrid approach. *Expert Systems with Applications*, *202*. https://doi.org/10.1016/j.eswa.2022.117370

Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. *Decision Analytics Journal*, *2*. https://doi.org/10.1016/j.dajour.2021.100015

Díaz, A., & Jareño, F. (2009). Explanatory factors of the inflation news impact on stock returns by sector: The Spanish case. *Research in International Business and Finance*, *23*(3), 349–368. https://doi.org/10.1016/j.ribaf.2008.12.001

Ding, D. K., Charoenwong, C., & Seetoh, R. (2004). Prospect theory, analyst forecasts, and stock returns. *Journal of Multinational Financial Management*, *14*(4–5), 425–442. https://doi.org/10.1016/j.mulfin.2004.03.005

Dogra, V., Singh, A., Verma, S., Alharbi, A., & Alosaimi, W. (2021a). Event study: Advanced machine learning and statistical technique for analyzing sustainability in banking stocks. *Mathematics*, *9*(24). https://doi.org/10.3390/math9243319

Dogra, V., Singh, A., Verma, S., Alharbi, A., & Alosaimi, W. (2021b). Event study: Advanced machine learning and statistical technique for analyzing sustainability in banking stocks. *Mathematics*, *9*(24), 1–18. https://doi.org/10.3390/math9243319

Engle, R. F., Hansen, M. K., Karagozoglu, A. K., & Lunde, A. (2021). News and Idiosyncratic Volatility: The Public Information Processing Hypothesis*. *Journal of Financial Econometrics*, *19*(1), 1–38. https://doi.org/10.1093/jjfinec/nbaa038

Fedorova, E., & Stepanov, V. (2023). The impact of innovation news coverage on illiquid stocks: the case of US market. *European Journal of Innovation Management*. https://doi.org/10.1108/ejim-07-2022-0387

Ftiti, Z., Ben Ameur, H., & Louhichi, W. (2021). Does non-fundamental news related to COVID-19 matter for stock returns? Evidence from Shanghai stock market. *Economic Modelling*, *99*. https://doi.org/10.1016/j.econmod.2021.03.003

Goodell, J. W., Kumar, S., Rao, P., & Verma, S. (2022). Emotions and stock market anomalies: A systematic review. *Journal of Behavioral and Experimental Finance*, *37*, 100722. https://doi.org/10.1016/j.jbef.2022.100722

Gregoriou, A., Healy, J. V., & Le, H. (2019). Prospect theory and stock returns: A seven factor pricing model. *Journal of Business Research*, *101*(April), 315–322. https://doi.org/10.1016/j.jbusres.2019.04.038

Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science, 17*, 26–32. https://doi.org/10.1016/j.procs.2013.05.005

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication*, *45*(4), 363–407. https://doi.org/10.1177/0021943608319388

Heston, S. L., & Ranjan Sinha, N. (2017). News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal*, *73*(3), 67–83. https://doi.org/10.2469/faj.v73.n3.3

Ho, K. Y., Shi, Y., & Zhang, Z. (2020). News and return volatility of Chinese bank stocks. *International Review of Economics and Finance*, *69*, 1095–1105. https://doi.org/10.1016/j.iref.2018.12.003

Iqbal, J., & Riaz, K. (2022). Predicting future financial performance of banks from management's tone in the textual disclosures. *Quality and Quantity*, *56*(4), 2691–2721. https://doi.org/10.1007/S11135-021-01216-5/TABLES/7

Jeon, Y., McCurdy, T. H., & Zhao, X. (2022). News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Journal of Financial Economics*, *145*(2), 1–17. https://doi.org/10.1016/j.jfineco.2021.08.002

Kabir Hassan, M. (2001). Is SAARC a viable economic block? Evidence from gravity model. In *Journal of Asian Economics* (Vol. 12).

Kim, S. J., Lee, L., & Wu, E. (2013). The impact of domestic and international monetary policy news on U.S. and German bank stocks. *International Finance Review*, *14*, 175–210. https://doi.org/10.1108/S1569-3767(2013)0000014010

King, R. G., & Levine, R. (1993). Finance and growth: Schumpeter might be right. *The Quarterly Journal of Economics*. http://qje.oxfordjournals.org/

Korenek, P., & Šimko, M. (2014). Sentiment analysis on microblog utilizing appraisal theory. *World Wide Web*, *17*(4), 847–867. https://doi.org/10.1007/s11280-013-0247-z

Levine, R. (1991). Stock Markets, Growth, and Tax Policy. *The Journal of Finance*, *46*(4), 1445–1465. https://doi.org/10.1111/j.1540-6261.1991.tb04625.x

Levine, R., & Zervos, S. (1998). Stock Markets, Banks, and Economic Growth. *The American Economic Association*, *88*(3), 537–558.

Li, Y., & Pan, Y. (2022). A novel ensemble deep learning model for stock prediction based on stock prices and news. *International Journal of Data Science and Analytics*, *13*(2), 139–149. https://doi.org/10.1007/s41060-021-00279-9

Lin, C. T., Wang, Y. K., Huang, P. L., Shi, Y., & Chang, Y. C. (2022). Spatial-temporal attention-based convolutional network with text and numerical information for stock price prediction. *Neural Computing and Applications*, *34*(17), 14387–14395. https://doi.org/10.1007/s00521-022-07234-0

Liu, Q., Tao, Z., Tse, Y., & Wang, C. (2022). Stock market prediction with deep learning: The case of China. *Finance Research Letters*, *46*. https://doi.org/10.1016/j.frl.2021.102209

Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications. *Environmental Modelling & Software*, *15*, 101–124.

Malik, F., Wang, F., & Naseem, M. A. (2017). ECONOMETRIC ESTIMATION OF BANKING STOCKS. *Source: The Journal of Developing Areas*, *51*(4), 207–237. https://doi.org/10.2307/26416972

Martin, J. R. (1995). Interpersonal meaning, persuasion and public discourse: Packing semiotic punch. *Australian Journal of Linguistics*, *15*(1), 33–67. https://doi.org/10.1080/07268609508599515

Martin, J. R., & White, P. (2005). The Language of Evaluation. In *The Language of Evaluation*. Palgrave Macmillan UK. https://doi.org/10.1057/9780230511910_1

Mittal, A., & Garg, A. K. (2021). Bank stocks inform higher growth—A System GMM analysis of ten emerging markets in Asia. *Quarterly Review of Economics and Finance*, *79*, 210–220. https://doi.org/10.1016/j.qref.2020.06.002

Mohsin, H. M., & Qayyum, A. (2005). *The Integration of Financial Markets: Empirical Evidence from South Asian Countries*.

Najaf, K., Joshipura, M., & Alshater, M. M. (2023). War build-up and stock returns: evidence from Russian and Ukrainian stock markets. *Journal of Risk Finance*, *24*(3), 354–370. https://doi.org/10.1108/JRF-05-2022-0107

Nemes, L., & Kiss, A. (2021). Prediction of stock values changes using sentiment analysis of stock news headlines. *Journal of Information and Telecommunication*, *5*(3), 375–394. https://doi.org/10.1080/24751839.2021.1874252

Nguyen, T. (2011). US macroeconomic news spillover effects on Vietnamese stock market. *Journal of Risk Finance*, *12*(5), 389–399. https://doi.org/10.1108/15265941111176127

Nyakurukwa, K., & Seetharam, Y. (2023). Can textual sentiment partially explain differences in the prices of dual-listed stocks? *Finance Research Letters*, 104529. https://doi.org/10.1016/j.frl.2023.104529

Pan, Z. (2019). A Review of Prospect Theory. *Journal of Human Resource and Sustainability Studies*, *07*(01), 98–107. https://doi.org/10.4236/jhrss.2019.71007

Peia, O., & Roszbach, K. (2015). Finance and growth: Time series evidence on causality. *Journal of Financial Stability*.

Samarasinghe, A. (2023). Stock market liquidity and bank stability. *Pacific Basin Finance Journal*, *79*. https://doi.org/10.1016/j.pacfin.2023.102028

Samarasinghe, A., & Uylangco, K. (2021). Stock market liquidity and traditional sources of bank business. *Accounting and Finance*. https://doi.org/10.1111/acfi.12883

Seng, J. L., & Yang, H. F. (2017). The association between stock price volatility and financial news – a sentiment analysis approach. *Kybernetes*, *46*(8), 1341–1365. https://doi.org/10.1108/K-11-2016-0307

Seng, J. L., Yang, P. H., & Yang, H. F. (2017). Initial public offering and financial news. *Journal of Information and Telecommunication*, *1*(3), 259–272. https://doi.org/10.1080/24751839.2017.1347762

Sensarma, R., & Jayadev, M. (2009). Are bank stocks sensitive to risk management? *Journal of Risk Finance*, *10*(1), 7–22. https://doi.org/10.1108/15265940910924463

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing Journal*, *90*. https://doi.org/10.1016/j.asoc.2020.106181

Sharma, S., Bhardwaj, I., & Kishore, K. (2022). Capturing the impact of accounting and regulatory variables on stock prices of banks – an empirical study of Indian banks in panel data modeling. *Asian Journal of Accounting Research*. https://doi.org/10.1108/AJAR-11-2020-0110

Shen, L. (2013). Financial dev, reforms and growth-main. *Economics Letters*, 215–219.

Sidhu, A. V., Rastogi, S., Gupte, R., & Bhimavarapu, V. M. (2022). Impact of Liquidity Coverage Ratio on Performance of Select Indian Banks. *Journal of Risk and Financial Management*, *15*(5). https://doi.org/10.3390/jrfm15050226

Stiglitz, J., Salvador, S., & Salvador, E. (1998). *The Role of the Financial System in Development (Speech) Seite 1 The Role of the Financial System in Development Senior Vice President and Chief Economist The World Bank*. http://www.worldbank.org/html/extdr/extme/jssp062998.htm

Tam Hoang, D. (2014). *Charles University in Prague Sentiment Analysis: Polarity Dataset*. http://www.imdb.com/reviews/156/15606.html.

Tauni, M. Z., Fang, H. X., & Iqbal, A. (2016). Information sources and trading behavior: does investor personality matter? *Qualitative Research in Financial Markets*, *8*(2), 94–117. https://doi.org/10.1108/QRFM-08-2015-0031

Uang, J.-Y., Citron, D. B., Sudarsanam, S., Taffler, R. J., Uang, J.-Y., Citron, D. B., Sudarsanam, S., & Taffler, R. J. (2006). Management Going-concern Disclosures: Impact of Corporate Governance and Auditor Reputation. In *European Financial Management* (Vol. 12, Issue 5).

Wojarnik, G. (2021a). Sentiment analysis as a factor included in the forecasts of price changes in the stock exchange. *Procedia Computer Science*, *192*, 3176–3183. https://doi.org/10.1016/j.procs.2021.09.090

Wojarnik, G. (2021b). Sentiment analysis as a factor included in the forecasts of price changes in the stock exchange. *Procedia Computer Science*, *192*, 3176–3183. https://doi.org/10.1016/j.procs.2021.09.090

Yao, X., Wu, D., Li, Z., & Xu, H. (2023). On the prediction of stock price crash risk using textual sentiment of management statement. *China Finance Review International*. https://doi.org/10.1108/CFRI-12-2022-0250

Yuan, H., Tang, Y., Xu, W., & Lau, R. Y. K. (2020). Exploring the influence of multimodal social media data on stock performance: an empirical perspective and analysis. *Internet Research*, *31*(3), 871–891. https://doi.org/10.1108/INTR-11-2019-0461

Zhang, M., Yang, J., Wan, M., Zhang, X., & Zhou, J. (2022). Predicting long-term stock movements with fused textual features of Chinese research reports. *Expert Systems with Applications*, *210*. https://doi.org/10.1016/j.eswa.2022.118312