IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

Predicting Stock Market Trends with Deep Learning

Mayur Selokar, Prof. Vijay Rakhade, Prof. Lowlesh Yadav

Student, Department of Computer Science & Engineering, Shri Sai College of Engineering & Technology, Chandrapur, India

Assistant Professor, Department of Computer Science & Engineering, Shri Sai College of Engineering & Technology, Chandrapur, India

Assistant Professor, Department of Computer Science & Engineering, Shri Sai College of Engineering & Technology, Chandrapur, India

ABSTRACT: The stock market is a highly studied field, and accurately predicting its movements is crucial today. However, it's a tough task requiring in-depth analysis of data patterns. Advanced statistical models and artificial intelligence algorithms are essential for this task. Machine learning and deep learning algorithms, in particular, offer promising results with minimal error margins. Among them, the Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) stand out, extensively used for stock market price prediction. These models predict future data based on recent trends, recursively updating as new data arrives. Efforts to enhance prediction accuracy using deep learning have shown significant progress. For instance, the ANN model achieved 97.66% accuracy, while the CNN model reached 98.92%. The CNN model employed 2-D histograms generated from quantified data within specific timeframes for prediction, a novel approach in this domain. To assess real-world applicability, the models were tested during the recent COVID-19 pandemic, Russia Ukrain war which caused significant market fluctuations. Remarkably, they yielded a 91% accuracy rate under these challenging conditions.

KEYWORD: Stock market, Machine learning, Prediction, Convolutional Neural Network (CNN) Artificial Neural Network (ANN)

I. INTRODUCTION

The stock market serves as the hub for buying, selling, and trading shares of publicly held companies, operating through regulated exchanges and over-the-counter markets. Given its dynamic and unpredictable nature, predicting stock market trends has become increasingly vital. With advancements in computational power, the accuracy and speed of stock market prediction have significantly improved, thanks to the integration of Artificial Intelligence and machine learning models.

In India, stock market activities predominantly occur on the Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE). While both follow similar trading protocols and settlement processes, BSE boasts over 5000 listed firms compared to NSE's 1600. However, the majority of trading volume and market capitalization is concentrated within approximately 500 firms on BSE. Both exchanges fiercely compete for order flow, promoting market efficiency and innovation, with arbitrageurs ensuring price parity between them.

The stock market serves as a barometer of economic sentiment and can influence Gross Domestic Product (GDP), measuring a nation's total economic output. Fluctuations in the stock market often mirror changes in consumer sentiment, impacting spending behavior and consequently, GDP growth. India's economy, heavily reliant on consumption, sees private spending accounting for a significant portion of GDP.

Machine learning algorithms offer valuable insights into economic trends, aiding in informed financial decision-making. This study aims to provide accurate stock market predictions, even amidst unprecedented events like the COVID-19 pandemic. The research is structured into several sections: exploring related work and the experimental dataset, introducing the proposed prediction model, detailing the system configuration, presenting empirical and graphical results, comparing the CNN method's advantages over traditional ANN methods, and concluding with final remarks.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

II. METHODOLOGY

Several researchers have tackled the challenge of predicting stock market trends, employing diverse methodologies. Jayanth Balaji employed deep learning techniques, testing 14 different methods to forecast stock prices. Similarly, Tsong Wuu Lin utilized Artificial Neural Networks (ANN) to optimize profitability. Autoregressive models, renowned for their accuracy in time series analysis, have also been applied to stock market prediction. Additionally, sentiment analysis, particularly through social media analytics, offers valuable insights into market sentiment. Techniques like ARIMA and deep learning models such as CNN and LSTM have been leveraged for sentiment analysis and forecasting. Despite the effectiveness of traditional models like Boosted Decision Trees and ELSTM, they struggle during volatile market conditions such as the COVID-19 pandemic. In contrast, Convolutional Neural Networks (CNNs) have emerged as a powerful tool for stock market prediction. While CNNs are traditionally used for image processing, they can effectively analyze time series data as well. By representing stock market data as 2-D histograms, CNNs demonstrate enhanced accuracy in prediction, even amid market fluctuations.

Building upon this research landscape, this study focuses on employing deep learning methods to predict stock market prices using the National Stock Exchange (NSE) dataset. Two approaches are explored: one utilizing backpropagation on a feed-forward neural network (ANN), and the other employing a convolutional neural network (CNN) model. The CNN model processes 2-D histograms generated from stock market data, offering a novel perspective and superior prediction accuracy.

The experimental dataset comprises NSE stock market data, specifically the NIFTY price index, spanning from April 2008 to April 2018. Additionally, recent data from November 2019 to August 2020 is analyzed as a case study. The dataset includes features such as 'Date,' 'Open,' 'Close,' 'High,' and 'Low.' By visualizing these features and generating 2-D histograms, patterns in the stock market data are identified to inform prediction models.

Two models are proposed:

1. ANN with Backpropagation Algorithm:

The backpropagation algorithm is a cornerstone of training artificial neural networks (ANNs), particularly in the context of feedforward neural networks like multi-layer perceptrons (MLPs). It operates through two main phases: the forward pass and the backward pass. During the forward pass, input data is fed into the network, and the activations of each neuron are computed layer by layer until the output is generated. This involves applying an activation function to the weighted sum of inputs for each neuron. In the backward pass, the error between the predicted output and the actual output is calculated using a chosen loss function, such as mean squared error. This error is then propagated backward through the network, layer by layer, to compute gradients of the loss function with respect to the network's weights. These gradients are utilized to update the weights of the network, typically using optimization algorithms like stochastic gradient descent (SGD). The learning rate parameter determines the magnitude of these weight updates. The training process involves iteratively feeding training data, performing forward and backward passes, and updating weights until convergence or a stopping criterion is met.

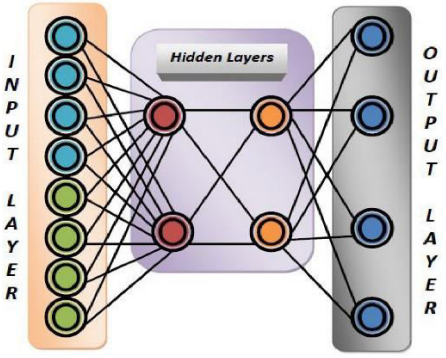

Figure 1: ANN model used in this prediction

- Designed as a deep feed-forward neural network.
- Consisted of eight input neurons representing the features of the previous two days.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

- Utilized two hidden layers, each with two neurons, and an output layer with four nodes.
- Activation function: Sigmoid.
- Trained with 2000 input combinations for over 10,000 epochs with a learning rate of 0.1.

Regularization techniques, such as L2 regularization (weight decay), can be applied to the backpropagation algorithm to mitigate overfitting by adding a regularization term to the loss function. Various hyperparameters need to be set, including the learning rate, batch size, number of epochs, and the architecture of the neural network (e.g., number of layers, number of neurons per layer, activation functions). Backpropagation finds applications in diverse domains, including classification, regression, image recognition, natural language processing, financial prediction, and control systems. It is crucial for training deep neural networks, enabling the learning of hierarchical representations of data and pattern recognition tasks. Overall, backpropagation has significantly advanced the field of neural networks and is a fundamental component of modern machine learning systems.

2. CNN with 2-D Histograms:

In the context of stock market prediction, the use of convolutional neural networks (CNNs) with 2-D histograms represents a novel approach to analyzing time-series data. Typically, CNNs are renowned for their effectiveness in tasks such as image recognition and computer vision, where they excel at capturing spatial patterns and hierarchies within data.

In the case of stock market prediction, the use of 2-D histograms involves representing the relationship between consecutive days' stock market features (such as open, close, high, and low prices) as a 2-dimensional grid. Each cell in the grid corresponds to a specific combination of feature values from two consecutive days. By binning these combinations into a grid and visualizing them using a heatmap or grayscale representation, the density of price movements and patterns over time can be captured.

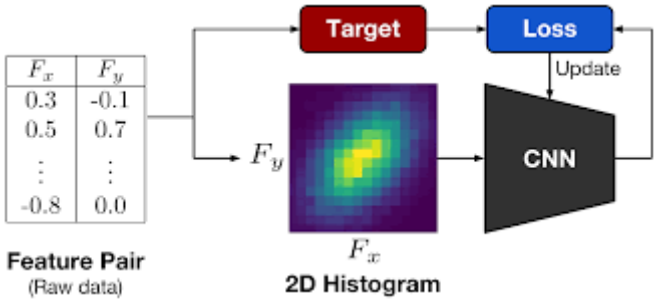

Figure 2: CNN model used in this study

- Utilized 2-D histograms generated from the stock market data as input.
- Trained with only 1000 input combinations and about 150 epochs.
- Utilized the Adam optimizer with a learning rate of 0.01 and Mean Squared Error (MSE) as the cost function.
- Achieved faster and efficient training compared to the ANN model.

The CNN architecture is then applied to learn features directly from these 2-D histograms. The network typically consists of convolutional layers followed by pooling layers, which enable the network to automatically extract relevant spatial patterns and features from the input data. The final layers of the CNN may include fully connected layers for classification or regression tasks, depending on the specific prediction problem.

During training, the CNN learns to identify patterns and relationships within the 2-D histograms that are indicative of future stock market movements. The training process involves optimizing the network's parameters (e.g., weights and biases) using techniques such as gradient descent and backpropagation.

One advantage of this approach is its ability to capture complex, non-linear relationships within the stock market data, which may be challenging to model using traditional methods. Additionally, by directly processing the 2-D histograms, the CNN can leverage spatial information and capture patterns that may not be apparent in the raw time-series data alone.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

Experimental Dataset:

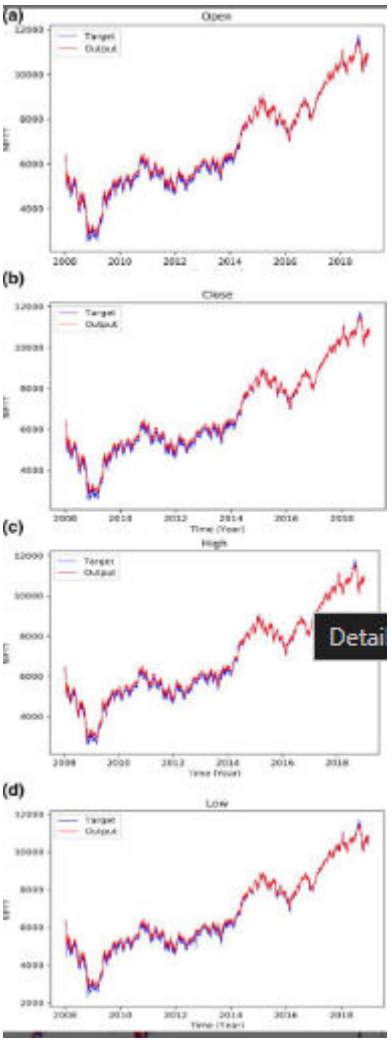

Figure 3: The prediction of all the features based on the ANN model

- The NSE stock market dataset, specifically the NIFTY price index, ranging from April 2008 to April 2018, was utilized.
- Additionally, recent data from November 2019 to August 2020 was considered as a case study.
- The dataset contains features such as 'Date,' 'Open,' 'Close,' 'High,' and 'Low.'

Quantisation Table:

The quantization table serves as a crucial evaluation tool for assessing the accuracy of the convolutional neural network (CNN) model in predicting stock market indices. In this study, the dataset was meticulously segmented into 15 distinct segments to conduct a comprehensive analysis of the CNN model's performance. Each segment represents a specific time frame within the dataset, allowing for a detailed examination of the model's predictive capabilities across different market conditions.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

Stock market data	'High' price index (Rs)		Accuracy (%)	'Low' price index (Rs)		Accuracy (%)
	Actual value	Predicted value		Actual value	Predicted value	
(Jan-2015)–(Oct-2015)	9119.2	9130.92	99.87147995	7539.5	7524.16	99.79653823
(Nov-2015)–(Sep-2016)	8968.7	9059.88	98.98335322	6825.8	6828.34	99.96278824
(Oct-2016)–(Apr-2017)	9367.15	9623.66	97.26160038	7893.8	7737.58	98.02097849
(May-2017)–(Feb-2018)	11,171.55	11,180.32	99.92149702	9269.9	9243.23	99.71229463
(Mar-2018)–(Dec-2018)	11,760.2	11,758.02	99.98146290	9951.9	9904.56	99.52431194

Figure 4: Accuracy determination by quantisation of the dataset

Within each segment, the actual and predicted 'High' and 'Low' values of the stock market indices were meticulously compared. This comparison enabled the calculation of the prediction accuracy for each segment, providing insights into the model's performance under various market dynamics. By considering both the 'High' and 'Low' values, the evaluation accounted for the inherent volatility and range of fluctuations present in stock market data, ensuring a comprehensive assessment of the model's predictive accuracy.

The results of this evaluation revealed an impressive overall accuracy of 99.34% for the CNN model. This high level of accuracy underscores the efficacy of the CNN approach in accurately forecasting stock market indices, even amidst the complexity and variability inherent in financial markets. Moreover, by quantifying the model's performance across multiple segments, this analysis provides valuable insights into the robustness and generalizability of the CNN model, demonstrating its ability to maintain high levels of accuracy across diverse market conditions and time periods.

III. CONCLUSION

The conclusion of the study highlights two distinct approaches to predict stock market indices and prices, each with its advantages and considerations:

1. Feed-forward Neural Network (FNN):

- Utilized the backpropagation algorithm for training.
- Provided fundamental insights into prediction trends.
- Achieved satisfactory results with an average accuracy of 97.66%.
- Required extensive training data and epochs, and suffered from overfitting, mitigated using regularization.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205142 |

2. Convolutional Neural Network (CNN) with 2-D Histograms:

- Offered a cleaner approach with better results on the given dataset.
- Utilized greyscale 2-D histograms for prediction, reducing training data and time.
- Segmented the dataset into 15 segments, enhancing prediction accuracy.
- Achieved an average prediction accuracy of 98.92%, surpassing the FNN model.

Both approaches have their pros and cons. The FNN model provides fundamental insights and satisfactory results but demands extensive training data and time. On the other hand, the CNN model with 2-D histograms offers better accuracy with reduced training requirements, albeit with an overhead in generating synthetic images.

Efficient prediction of the stock market is crucial in today's world, benefiting stock market analysts in developing solutions to aid companies and economies by predicting future market patterns. The study's findings provide valuable tools for stock market analysis and decision-making..

REFERENCES

1. Lowlesh Yadav and Asha Ambhaikar, “IOHT based Tele-Healthcare Support System for Feasibility and perfor-mance analysis,” Journal of Electrical Systems, vol. 20, no. 3s, pp. 844–850, Apr. 2024, doi: 10.52783/jes.1382.
2. L. Yadav and A. Ambhaikar, "Feasibility and Deployment Challenges of Data Analysis in Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-5, doi: 10.1109/ICAIIHI57871.2023.10489389.
3. L. Yadav and A. Ambhaikar, "Approach Towards Development of Portable Multi-Model Tele-Healthcare System," 2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI), Raipur, India, 2023, pp. 1-6, doi: 10.1109/ICAIIHI57871.2023.10489468.
4. Lowlesh Yadav and Asha Ambhaikar, Exploring Portable Multi-Modal Telehealth Solutions: A Development Approach. International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), vol. 11, no. 10, pp. 873–879, Mar. 2024.11(10), 873–879, DOI: 10.13140/RG.2.2.15400.99846.
5. Lowlesh Yadav, Predictive Acknowledgement using TRE System to reduce cost and Bandwidth, March 2019. International Journal of Research in Electronics and Computer Engineering (IJRECE), VOL. 7 ISSUE 1 (JANUARY- MARCH 2019) ISSN: 2393-9028 (PRINT) | ISSN: 2348-2281 (ONLINE).
6. Yu Liu, Lei Wang, and Bing Li. "A Comprehensive Survey of Deep Learning for Image Recognition." IEEE Transactions on Neural Networks and Learning Systems (2019).
7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details