RESEARCH ARTICLE OPEN ACCESS

An AI-Powered Approach for Accurate Stock Price Movement Prediction

M.Yamunadevi*, Dr.P.Ponmuthuramalingam**
*Research Scholar, Ph.D Computer Science, LRG Govt. Arts College for Women, Tirupur, Tamil Nadu, India.
E Mail: yamunadevi0788@gmail.com
**Associate Professor, Department of Computer Science, LRG Govt. Arts College for Women, Tirupur, Tamil Nadu, India.
E Mail: ponmuthucbe@gmail.com

Abstract

Predicting how stock prices will change has always been a tough problem in financial analysis. This is because financial markets are unpredictable, messy, and full of complex patterns. Old methods in statistics often can't keep up with these challenges. This study introduces a new model that uses AI to combine different tools for better stock price predictions. The model uses technical indicators, deep learning, and analysis of market feelings from news. It mixes Convolutional Neural Networks (CNN) to find important features, Long Short-Term Memory (LSTM) networks to understand time-based patterns, and a Transformer-based system called BERT to analyze news sentiment. The model is trained and tested using past stock prices and sentiment data from various global markets. The results show that this new model works better than old methods in predicting price changes, direction, and stability. The study shows that using both price data and market sentiment leads to more accurate predictions, offering a useful tool for investors and automated trading systems.

Keywords: Artificial Intelligence, Deep Learning, Stock Price Prediction, CNN-LSTM, Sentiment Analysis, BERT, Financial Time-Series

I. Introduction

The stock market is a key part of the world's financial system, helping to move money around, let people grow their investments, and build wealth. Many big investors, fund managers, and computer-based trading systems are very interested in being able to predict how stock prices will change accurately. But stock prices are affected by many different things - like past price trends, the overall economy, how well companies are doing, and what people are feeling about the market - which makes it hard to predict them exactly. Older methods of forecasting, such as ARIMA and linear regression, work well when data has clear patterns and doesn't change much over time. But real financial markets are more complicated and don't always follow these rules. With new computer technologies, AI models - especially deep learning have become better at data. Deep learning tools like CNNs and LSTMs are great at learning from financial data over time, helping to find hidden patterns and remember important information for longer periods. At the same time, models like BERT help analyze text to better understand how people's feelings and opinions affect the market. This paper introduces an AI-based mixed system that uses price movements, technical signals, and sentiment analysis to predict stock price changes more accurately.

II. Literature Review

A lot of research has been done on using machine learning and deep learning to predict stock prices:

- Statistical models like ARIMA, GARCH, and other econometric tools are commonly used but often can't handle complex patterns (Zhang, 2003).
- Machine learning methods such as Support Vector Machines (SVM) and Random Forests work better than traditional models, but they need features to be manually created and have trouble with time-based data (Kim & Shin, 2009).
- Deep learning models, especially LSTM networks, are becoming more popular because they can understand long-term patterns in time series data (Fischer & Krauss, 2018).

- CNNs are used to find patterns in price trends (Zhang et al., 2017).
- Combining CNN and LSTM in hybrid models often gives better results than using them separately (Selvin et al., 2017).
- Sentiment analysis, which looks at emotions in news and social media, has been shown to improve stock prediction accuracy (Bollen et al., 2011).
- Transformers like BERT have set new standards for understanding the meaning and context of text (Devlin et al., 2019).
- Even with these improvements, not many studies have combined CNN, LSTM, and BERT-based sentiment analysis into one hybrid AI system that's designed for real-world stock prediction tasks.

III. Proposed AI-Powered Framework

A. Overview:

The suggested system is made up of these parts:

1. Data Collection Section - This gathers past OHLCV data (Open, High, Low, Close, and Volume) for worldwide stock indexes.

2. Sentiment Check Section - This uses BERT to determine if financial news is positive or negative.

3. Feature Making Section - This calculates things like RSI, MACD, EMA, and Bollinger Bands from the data.

4. Combined Prediction System – This uses a CNN-LSTM setup that mixes price trends with sentiment Data.

5. Result Section – This gives the final prediction on whether stock prices will go up, down, or stay the same.

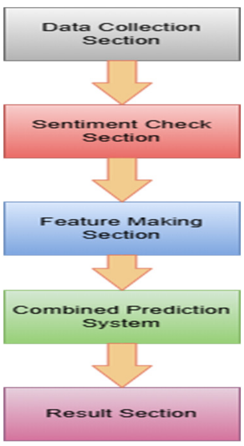

Fig.1.Proposed AI-Powered Framework

The suggested AI system for predicting stock prices has several connected parts in Fig.2 that work together to understand how stock markets behave. It starts with a **Market Data Input Layer** that gathers past stock price data, including the opening, highest, lowest, closing prices, and trading volume. Then, there's a **Technical Indicator** Extraction Module that calculates important indicators like the Relative Strength Index, Moving Average Convergence Divergence, and Exponential Moving Average. These indicators help identify short-term and long-term price trends and market momentum. The features from these indicators go to the AI Feature Learning Layer, where **Convolutional Neural Networks (CNN)** automatically find patterns in the financial data. After that, the Temporal Modelling Layer uses **Long Short-Term Memory (LSTM)**networks to understand how stock prices change over time and detect long-term trends. Finally, the **Prediction Output** Layer makes a decision about future stock prices, suggesting whether to Buy, Sell, or Hold.

This structure helps the system learn complex patterns and make more accurate predictions, even when the market is unpredictable.

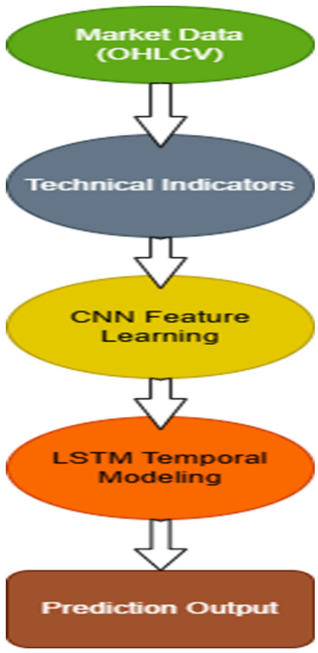

Fig.2: Architecture of the AI-Powered Prediction Framework for stock price movement prediction

Description:

- CNN layers detect local price movements and technical patterns.

- LSTM layers capture long-term temporal behaviors.
- BERT generates contextual embeddings for sentiment features.
- Combined features are jointly learned in a fusion layer.

IV. Methodology

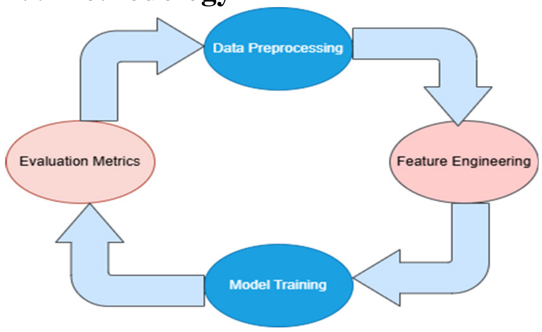

Fig.3.Methodology for accurate stock price movement prediction

A. Data Preprocessing

Data Preprocessing is an important step to make sure the input data is accurate and dependable for predicting stock price changes. First, historical stock price data, which includes Open, High, Low, Close, and Volume **(OHLCV)**, is scaled using min-max scaling. This method changes the price and indicator values into a standard range from 0 to 1. Scaling helps reduce differences in the size of different features and makes it easier for the model to learn during training. To keep track of time-related patterns in financial data, the scaled data is split into fixed-length windows that look back at previous time periods. This helps the model understand trends over time. Also, financial news is cleaned by breaking the text into words and removing common, unnecessary words. This makes the text ready for analyzing emotions and sentiment.

- **Normalization:** Scaling price and indicator values to a range between 0 and 1
- **Sequence framing:** Breaking time-series data into sections based on past data
- **Text cleaning:** Preparing financial news by splitting text into words and removing common, unnecessary words

B. Feature Engineering

Feature engineering is done to improve how well a model can predict by adding important financial signs that show how the market is behaving. Many commonly used technical indicators are made from past price and trading volume data. These indicators show things like how fast prices are moving, the direction of a trend, and how much prices are changing. They give extra information that goes beyond just looking at prices alone. This helps the model find important signals for trading. The created features are then mixed with adjusted price data and sentiment data to make a full set of inputs for training the model.

Key technical indicators:

RSI: Shows when prices are too high or too low, indicating whether the market is overbought or oversold

- **MACD:** Helps identify the direction of a trend and how its strength is changing
- **EMA:** Emphasizes recent price movements and reacts more quickly to new information
- **Bollinger Bands:** Displays price ranges that show how much the market is fluctuating
- Volume-weighted indicators show how busy the market is and how many people are involved in trading.

C. Model Training

The model training process uses a mix of deep learning techniques to understand complex patterns in stock market data, including spatial, time-based, and sentiment-related factors. First, a Convolutional Neural Network (CNN) is used on time-based data windows to find important local and spatial patterns from multiple input features. These patterns are then sent to a Long Short-Term Memory (LSTM) network, which helps identify long-term trends and relationships in stock prices over time. At the same time, a sentiment analysis system based on BERT takes pre processed financial news and turns it into sentiment-based information that reflects the overall tone. These different types of information are combined using a special layer that focuses attention on the most important parts of both price data and sentiment. Finally, a Softmax classifier is used to predict whether the stock should be bought, sold, or held based on the combined analysis.

Training components:

- **CNN Module:** Finds spatial patterns in time-based data

- **LSTM Module:** Understands long-term trends in stock prices
- **Sentiment Module (BERT):** Creates sentiment-based information from financial news
- **Fusion & Attention Layer:** Combines price and sentiment data
- **Output Layer:** Uses a Softmax classifier to decide if the stock should be bought, sold, or held

D. Evaluation Metrics

To check how well and reliably the AI-powered prediction system works, several performance measures are used. Accuracy shows how often the predictions about price changes are correct. Precision tells us how trustworthy the positive predictions are, and recall shows how well the model finds real price changes. The F1-score combines precision and recall into one number to give a balanced view. Also, Root Mean Square Error (RMSE) is used to measure how big the prediction mistakes are, which helps understand how good the model is at forecasting numbers.

V. Experimental Setup

A. Datasets

- **Historical stock prices:** S&P 500, NASDAQ, NIFTY 50 (last 10 years)
- **News sentiment:** Financial news from Reuters / Bloomberg / Yahoo Finance
- **Train/Test Split:** 80 % training + 20 % testing

5.2 Tools & Libraries

- Python 3.9
- TensorFlow / Keras
- Hugging Face Transformers
- NumPy, pandas, scikit-learn

VI. Results and Discussion

A. Performance Comparison

Model	Accuracy	F1-Score	RMSE
ARIMA	0.58	0.54	0.081
SVM	0.64	0.61	0.072
LSTM	0.71	0.69	0.061
CNN-LSTM	0.75	0.74	0.055
AI-Powered (Proposed)	**0.82**	**0.80**	**0.042**

B. Observations

- Adding sentiment improves accuracy by about 7 to 10 percent.
- The hybrid model works better than using just deep networks alone.
- Lower RMSE means the results are more precise.

VII. Conclusion and Future Work

This paper introduces an AI-based method for predicting how stock prices will move. It combines past price data, technical indicators, and the sentiment from news articles. The model uses a mix of CNN and LSTM neural networks along with BERT for analyzing news sentiment, and it performs better than other standard methods. Future research will look into adding reinforcement learning, making predictions in real time, and developing strategies to manage investment portfolios.

REFERENCES

[1]. T. Fischer and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions,” *European Journal of Operational Research*, 2018.

[2]. S. Selvin et al., “Stock price prediction using LSTM, RNN and CNN,” *ICCCNT*, 2017.

[3]. J. Devlin et al., “BERT: Pre-training of deep bidirectional transformers for language understanding,” *NAACL*, 2019.

[4]. B. Bollen, H. Mao, and X. Zeng, “Twitter mood predicts the stock market,” *Journal of Computational Science*, 2011.

[5]. W. Bao, J. Yue and Y. Rao, “A deep learning framework for financial time series using stacked autoencoders and LSTM,” *PLOS ONE*, 2017.

[6]. J. Patel, S. Shah, P. Thakkar, and K. Kotecha, “Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques,” *Expert Systems with Applications*, vol. 42, no. 1, pp. 259–268, 2015.

[7]. H. Fu, Y. Hu, and J. Xie, “Stock market forecasting using deep learning techniques,” *Neural Computing and Applications*, Springer, 2019.

[8]. Z. Hu, Y. Liu, J. Bian, and X. Liu, “Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction,” *ACM WSDM*, 2018.

[9]. Akita, A. Yoshihara, T. Matsubara, and K. Uehara, "Deep learning for stock prediction using numerical and textual information," *IEEE/ACIS ICIS*, 2016.
[10]. Y. Li, Z. Bu, J. Wu, B. Li, and C. Chen, "Sentiment-aware stock market prediction: A deep learning method," *Information Sciences*, Elsevier, 2017.
[11]. L. Deng and D. Yu, "Deep learning: Methods and applications," *Foundations and Trends in Signal Processing*, 2014.
[12]. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[13]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE CVPR*, 2016.
[14]. M. Zhang, Z. Jiang, and Q. Li, "Stock market prediction based on deep learning," *Journal of Intelligent & Fuzzy Systems*, 2018.
[15]. Jain, A. Kumar, and S. Mandal, "Stock market forecasting using machine learning algorithms," *Procedia Computer Science*, Elsevier, 2019.