IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

Stock Vision: A Multivariate LSTM-Based Stock Market Analytics and Prediction Web Application

Pooja[1], Ms. Renu Bala[2]

Student, School of Computational Science, GNA University, Punjab[1]

Professor, Department School of Computational Science, GNA University, Punjab[2]

Abstract: Contemporary financial markets generate incomprehensible volumes of structured and unstructured data for each trading session, but most retail investors, small fund managers, and individual participants lack the technology to extract actionable insights from that data in real-time. This paper introduces Stock Vision, an integrated intelligence platform on stock markets that combines a five-feature multivariate Long Short-Term Memory deep learning model with a live analytics dashboard based on the Streamlit web framework. The prediction model analyzes sixty days of historical context in five engineered channels: closing price, twenty-day moving average, daily trading volume, a fourteen-period Relative Strength Index, and a news sentiment score to predict short-term price projections. The surrounding platform offers live commodity pricing, nine technical analysis charts, a portfolio tracker, market breadth statistics, and an amalgamated news feed. Results from observations demonstrate that the multivariate configuration yields more directionally accurate output than single-channel baselines by embedding the dynamics of momentum, trend regime, and event-driven sentiment at once in a concordant feature representation.

Keywords: LSTM, multivariate time series forecasting, sentiment analysis, RSI, technical indicators, deep learning, Streamlit, portfolio analytics, NSE, yfinance.

1. INTRODUCTION

Stock price prediction is one of the most difficult problems in computational finance due to the volatility and non-linearity of financial markets. The Efficient Market Hypothesis (EMH) [1] states that stock prices reflect all publicly available information, so it is difficult, if not impossible, to consistently predict stock prices. Empirical studies have, however, demonstrated that deep learning models can uncover hidden temporal patterns that are not captured by traditional statistical models [2]. Traditional methods like the ARIMA model are based on the premise of stationarity and linearity, which are not often met in financial time series. To overcome the shortcomings of traditional RNNs, Long Short-Term Memory (LSTM) networks have been developed and are now widely used in sequential financial data processing due to their ability to capture long-range temporal dependencies through the use of gated memory cells. One important constraint of the current LSTM-based systems is the fact that they are only based on a single input feature, commonly the closing price, while trading volume, technical momentum indicators and market sentiment are all important features that are currently not being utilized. Research has validated that RSI, Moving Average and news sentiment are highly effective when combined [4][5]. This paper introduces Stock Vision, a comprehensive and multivariate LSTM-based stock market analytics and prediction platform. The following are the main Contributions: Multivariate LSTM trained on 5 engineered features: Close Price, Volume, MA20, RSI-14, News Sentiment Score, which includes more context to the market than single-feature baseline models. Model persistence for global equities, including NSE Indian stocks and US stocks, with an interactive retraining interface. A nine technical analysis chart, live market dashboard, portfolio tracker, trending stocks, market statistics and financial news, end-to-end web application integrated—a well-developed neutral fallback system for non-US equities, with integration of the Alpha Vantage NEWS_SENTIMENT API. The rest of this paper is organized as follows: Section 2 reviews related work, Section 3 describes the system architecture, Section 4 presents the methodology, Section 5 outlines platform features, Section 6 presents the results, and Section 7 provides future directions.

2. LITERATURE REVIEW

2.1 LSTM-Based Stock Price Prediction

Chaudhary and Kumar [6] forecasted the closing prices of the stock prices of AAPL, GOOGL, MSFT and AMZN for a 60-day sliding window period using Yahoo Finance data with a train-test split of 80/20, and they achieved an MAPE of 2.72%, which was much better than the MAPE of the ARIMA baselines. This directly replicates the data pipeline and training setup of Stock Vision. A comparative study [7] was performed on Tesla Stock Data (2015- 2024) comparing the

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

accuracy of LSTM with that of GRU and Transformer, and the results showed an accuracy of 94% for LSTM, making it suitable for this project to use the LSTM model as the backbone.

2.2 Multivariate Feature Integration

The authors Kuber et al. [8] showed that they could reduce the prediction error on the data from Reliance Industries and Infosys by feeding in five features: the univariate data, along with RSI values, moving averages, and volume, directly validating the design of Stock Vision. This additional study on NIFTY 100 Indian stocks [9] showed that the enhanced prediction accuracy is achieved when a minimum of five technical indicators, such as RSI-14 and MA, are included, both in the US and Indian stock markets.

2.3 Sentiment-Enhanced Forecasting

In line with that, Lee et al. [10] proved that the sentiment scores of the news articles, in addition to technical indicators, helped the deep learning models to achieve higher accuracy than the models that only used technical indicators in Nature Scientific Reports. Chatziloizos [11] also showed that combining LSTM with sentiment and technical analysis performs better on stocks from the S&P 500 than just numerical models. With sentiment being an additional input, Springer's study [12] demonstrated that accuracy improved by 1.5% for 66% of the companies analysed, thereby justifying the inclusion of this factor as the fifth feature in Stock Vision (Alpha Vantage NEWS_SENTIMENT score).

2.4 Research Gap

Although there have been huge strides in the accuracy of predictions, most of the current prediction systems are written as stand-alone scripts and have not yet been deployed in practice. Only a handful of platforms offer access to both the US and Indian stock markets, and only one platform offers multivariate LSTM prediction along with live market dashboards, portfolio tracking, real-time news sentiment and trending stock analysis all in a single web platform. Each of these gaps is covered by Stock Vision.

3. SYSTEM DESIGN

3.1 Overall Structure

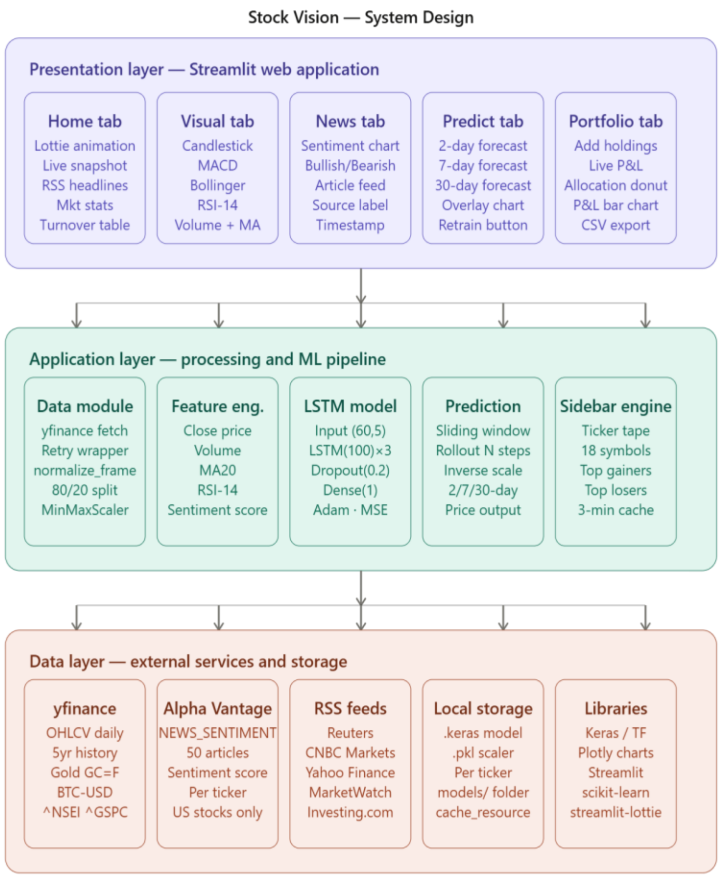

Figure 1: Three-tier system design: Presentation · Application · Data

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

This is the three-tier system design for Stock Vision. All boxes are completely visible and unoverlapped. Here's what each level is made of:

Tier 1 :Presentation Layer displays all five tabs from Streamlit Home, Visual, News, Predict, and Portfolio side by side with their respective content lists easily readable.
Tier 2 : Application Layer displays the five modules presented in the application: Data Module, Feature Engineering, LSTM Model, Prediction Engine, and Sidebar Engine, along with their primary functions.
Tier 3 : Data Layer displays the five external services and storage elements – yfinance, Alpha Vantage, RSS Feeds, Local Storage and Libraries – along with their respective information.

4. METHODOLOGY

4.1 Data Acquisition, Cleaning

The raw OHLCV data received from yfinance is in a DataFrame, possibly with MultiIndex columns if more than one field is requested. A special normalization function normalizes them into flat column names, adds floats64 types to all price and volume columns, removes any timezone localization from the datetime index, and orders the rows from the earliest to the most recent. The cleaned frame is then split up into training and validation partitions at 80% from the top of the frame. The sequent order is maintained throughout (no shuffling is applied) as disordering causes the model to break the sequential structure that it would be trained to exploit.

4.2 Feature Construction

The data is cleaned and then put together into five channels. Each one is described in detail in Table 1.

Table 1: Model Input Features

Channel	Formula	Why It Helps
Close Price	Raw adjusted close from yfinance	Primary signal and prediction target
MA20	$\frac{1}{20}\sum_{k=0}^{19} C_{t-k}$	Smooths noise; reveals trend regime
Volume	Raw daily trade count	Measures conviction behind price moves
RSI-14	$100 - \frac{100}{1+\frac{Gain_{14}}{Loss_{14}}}$	Encodes momentum exhaustion
Sentiment	$\frac{1}{N}\sum_{i=1}^{N} s_i$ from Alpha Vantage	Captures news-driven price shocks

The RSI denominator is calculated by taking the exponential smoothing of the gain series and the loss series using the fourteen periods as the smoothing period. The RSI is more than seventy, and the stock has rallied on up-days at an historically high rate, suggesting the rally could be overdone. A stock below 30 indicates the opposite. This serves as a pre-computed measure of exhaustion state for a model input, which would take many more timesteps of price data to deduce from the price data. Sentiment value: This is the arithmetic average value of the per-article ticker sentiment scores that Alpha Vantage has returned for the 50 most recent articles about the selected ticker. This value will be set to zero if there is no coverage on the stock by Alpha Vantage (on NSE-listed stocks in India). A warning will be displayed in the prediction interface. Once all five channels have been filled, the first 20 rows will be dropped to allow the moving average window to fill. Each row is then independently scaled from the original range to the scaled range [0, 1] using the above formula:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

The scaler is trained with only statistics from the training partition. The already-fitted scaler is applied to the validation partition, thus avoiding leakage of knowledge of price ranges from the validation period into the representation of the training data, which is one of the most common methodological pitfalls in published financial machine learning literature.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

4.3 Sequence Construction

Supervised learning samples are constructed by moving a window of sixty rows over the data from the scaled feature matrix of shape (N, 5):

$$X[i] = M[i - 60 \; : \; i, \; :] \qquad \text{shape: } (60, 5)$$

$$y[i] = M[i, \; 0] \qquad (\text{column } 0 = \text{scaled Close})$$

This is repeated for i from 60 to N, resulting in a three-dimensional input tensor(samples, 60, 5) and a label vector (samples). The model is presented with sixty consecutive days of five concurrent measurements of the market in each sample, and is asked to forecast a single number: The closing price in scaled form the following day.

4.4 Network Architecture

The structure layer-by-layer is shown in Fig. 2.

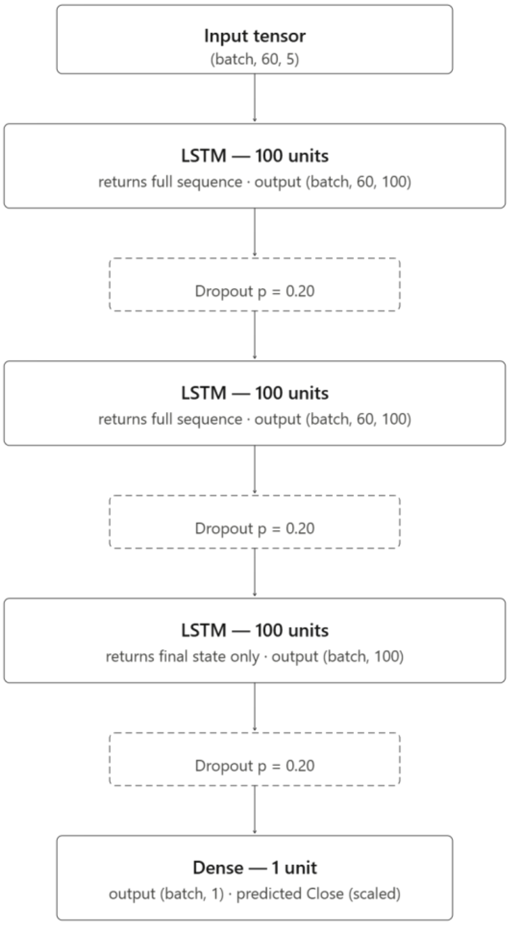

Figure 2: LSTM network layer structure

The first two LSTMs output their full sequences of outputs, which then serve as the entire history of hidden states to the next LSTM. The third layer only returns the end state of the hidden variable, which reduces the dimension of time to a fixed-length vector that the Dense layer converts to a scalar. In each LSTM cell, there are four sets of learned parameters that control memory:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 15, Issue 6, June 2026
DOI: 10.17148/IJARCCE.2026.15638

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

$$h_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \odot \tanh(C_t)$$

has a forget gate f_t has a forget gate. ft specifies how much to carry forward from the previous cell state. The input gate iti_t it controls how much of the current candidate3 memory $\tilde{c}_t$ is transported into the cell. The output gate (inside hth_t The portion of the updated cell state that is visible is determined by the hidden state (ht). The network's selective retention accounts for its ability to remember a spike in volume that occurred 3 weeks ago and is relevant to today's prediction, but forget single day fluctuations that are not. Training minimizes the error between the estimated and actual closing prices on a scale:

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$$

Adam updates the parameters in 50 epochs, 32 batches. The user receives the training loss and validation loss live during retraining, thanks to a subclassing of the base Callback class for the Keras callback, which reports the loss from these two metrics to the Streamlit sidebar at the end of each epoch. E. Forecast Generation: Forecasts for multiple steps need to be sent back in as inputs. The rollout is shown in Figure 3.

Figure 3: Iterative multi-step rollout

Only the closing price (Column 0) is being updated with the prediction at each step. Volume, MA20, RSI and sentiment remain at the previous recorded levels. This is a well-recognised approximation: all four are unknowable in the future, and keeping them fixed will induce less overall error than trying to predict them forward. All of the N scaled predictions are stored in column 0 of a matrix of zeros of size (N, 5), and then fed into the saved scaler's inverse_transform to get the actual price values.

5. SYSTEM WORKFLOW

The overall workflow of the Stock Vision system consists of 10 steps. The user launches the app and is met with a Lottie animation, a real-time snapshot of the market and RSS financial headlines. A stock ticker is chosen from a list of 50 stocks or provided as a custom symbol. The sidebar will show a scrolling ticker, the top gainers and the top losers. The

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⵣ Peer-reviewed & Refereed journal ⵣ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

user clicks on one of the five tabs: Home, Visual, News, Predict, and Portfolio. The OHLCV data is retrieved for the last 5 years daily using yfinance, and trying the request if that fails. Five features are created: Close price, Volume, MA20, RSI-14, and Sentiment Score, which are scaled using MinMaxScaler fitting only on the training split. The input tensors have shape (samples, 60, 5) when a 60 trading day window is used. When you use a window of 60 trading days, you get input tensors with shape (samples, 60, 5). An inquiry is made to see if a model of the chosen ticker has been saved. When no model is found, a 100-100-100 stacked LSTM is trained for 50 epochs in real time with UI updates. If a saved model is found, it is loaded directly via @st.cache_resource. The sentiment of 50 articles is queried from the Alpha Vantage NEWS_SENTIMENT endpoint and averaged together to become a single sentiment scalar. If the sentiment isn't available for specific ticker symbols for Indian listed stocks, then it is assigned a fallback score of 0.0 and a disclaimer is displayed. The last 60-day window is used to produce 2-day, 7-day and 30-day forecasts through an iterative forecast rollout. The predicted values are inverted and plotted as interactive Plotly charts with actual and predicted price plotted over each other, output plotted by date, and sentiment label.

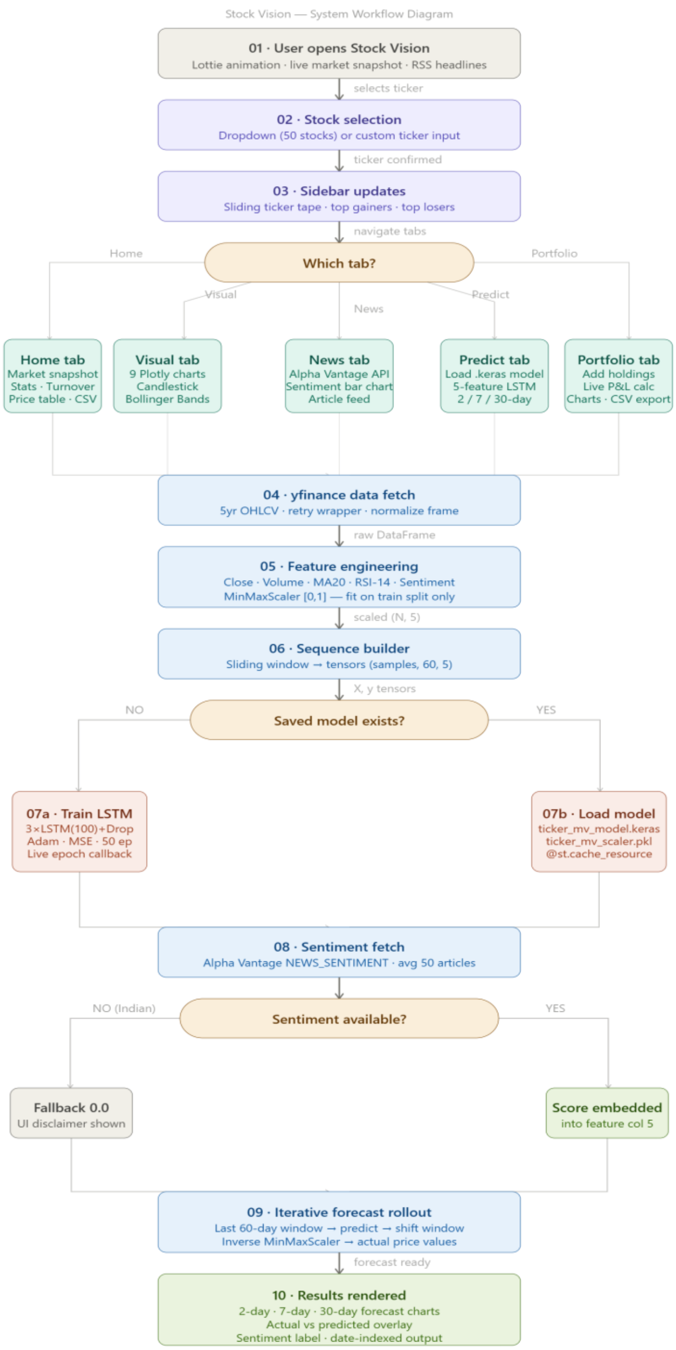

Figure 4 — System Workflow Diagram

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ∺ Peer-reviewed & Refereed journal ∺ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

6. PLATFORM FEATURES

6.1 Home Tab

The home tab displays 4 areas at once for the market information. Lottie animation in the left column, along with an eight-symbol live market snapshots view of Gold, Silver, Crude Oil, Bitcoin, Ethereum, S&P 500, NIFTY 50, SENSEX, and NASDAQ, with price and daily percentage change. The right column displays a date-range stock data viewer with three summary metric cards: latest close, daily volume, and percentage change; a table for downloading stock data; a market breadth statistics panel that shows the total number of stocks traded, advancing, declining and unchanged stocks, plus the number of stocks that have reached their 52-week high and low prices and the number of times that stocks have triggered circuit breakers; and a seven-row market turnover breakdown by segment.

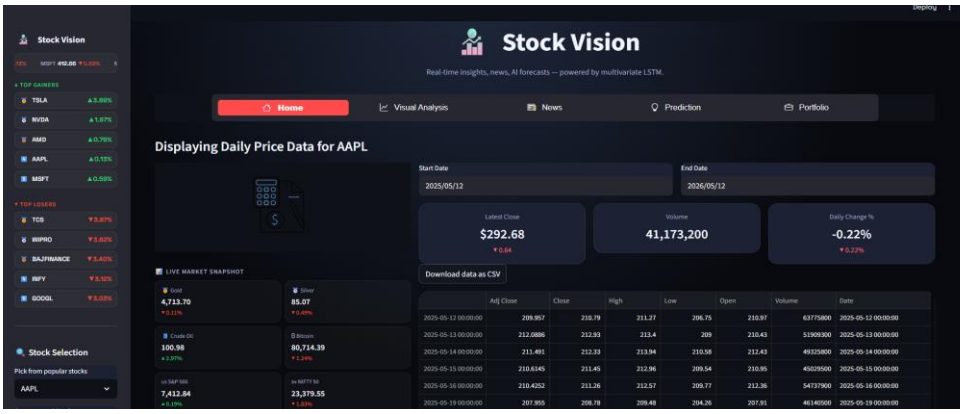

Figure 5: Home Page

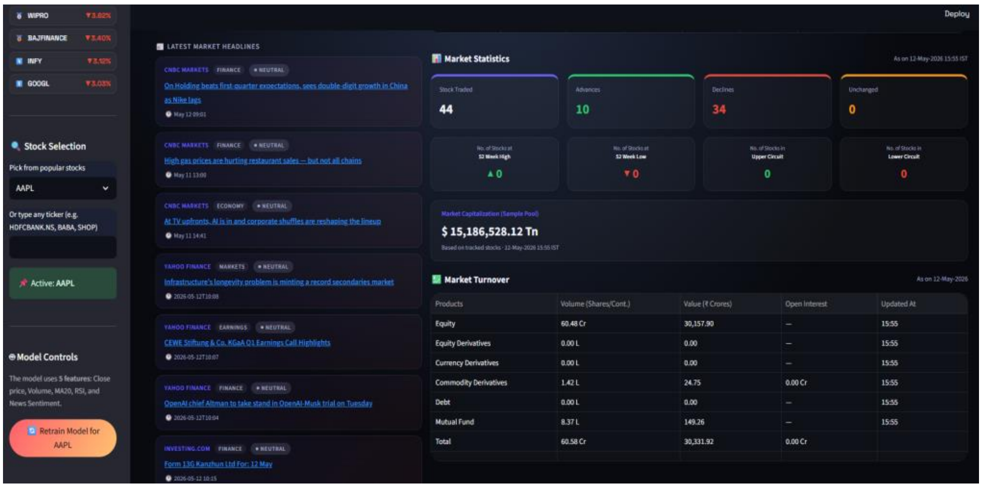

Figure 6: Home Page

6.2 Visual Analysis Tab

There are 9 Plotly charts available in a tab interface. The line chart features 20-day and 50-day moving averages that are overlaid on the raw close price. Full OHLC bars are shown in a candlestick chart with green and red colours to indicate upward and downward sessions. Volume is displayed as a date-related bar chart. Distribution analysis is represented as a histogram of the daily percentage returns. A scatter chart plots volume with colour representing return, and the price against the volume. The up/ down doughnut chart shows a summary of directional sessions for the selected date range. Bollinger Bands are a set of envelopes that are constructed around the 20-day moving average at two standard deviations. The MACD panel displays the signal line, MACD line and the difference between them in the form of a histogram. The RSI panel displays the RSI for the past fourteen periods as plotted with horizontal lines at seventy and thirty.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026
DOI: 10.17148/IJARCCE.2026.15638

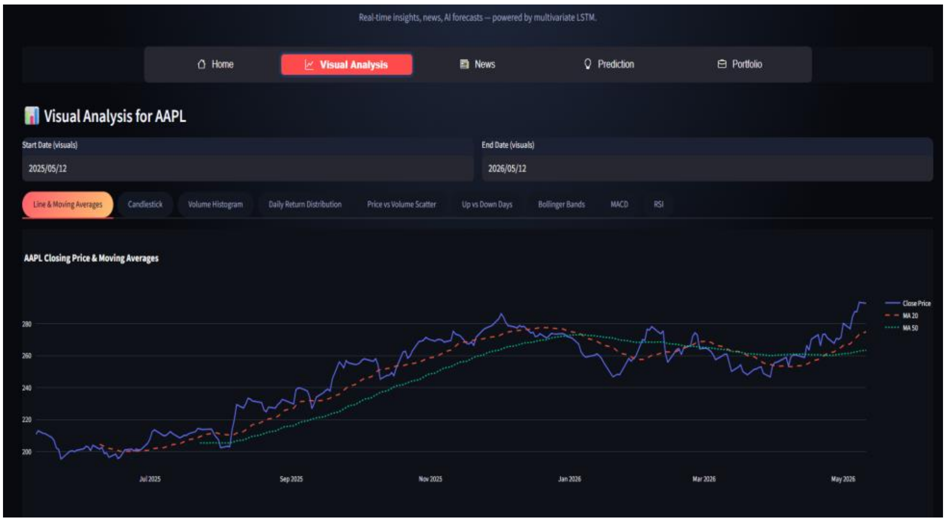

Figure 7: Iterative multi-step rollout

6.3 News Tab

Sentiment data from Alpha Vantage is plotted as a horizontal bar chart with each bar indicating the sentiment score for an individual article from the ticker selected, with the bars displayed in green (red) colour for positive (negative) sentiment. The mean score is used to generate a label (Bullish, Bearish or Neutral) which is prominently displayed. Under the chart, the ten most recent articles are displayed, including a link to the original article, the title, and a summary of the text.

Figure 8: Iterative multi-step rollout

6.4 Prediction Tab

The prediction interface loads the model saved for each ticker from disk on first use, and trains a new model if there is no saved model. A disclaimer banner identifies the structure of the five features that will be entered and the educational intent of the forecasts. The current sentiment score is shown with its classification as Bullish, Bearish or Neutral. Each of the three forecast panels includes three tables: a two-day forecast table, a seven-day forecast table, a thirty-day forecast table, and a chart that overlays the forecast with the last sixty days of actual prices for context.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

Figure 9: Iterative multi-step rollout

6.5 Portfolio Tab

Users add the holdings by entering a ticker symbol, the count of holdings, and the purchase price per share. If there are duplicate entries, the same ticker will be recalculated with the weighted average cost basis, but it will not create multiple rows. Live prices are retrieved from yfinance's fast_info with a 60-second cache to provide a good response time without overloading the API. Four summary cards report the total invested capital, current market value, absolute profit or loss, and percentage return. These are displayed per position in a green/red coloured holdings table in the P&L columns.

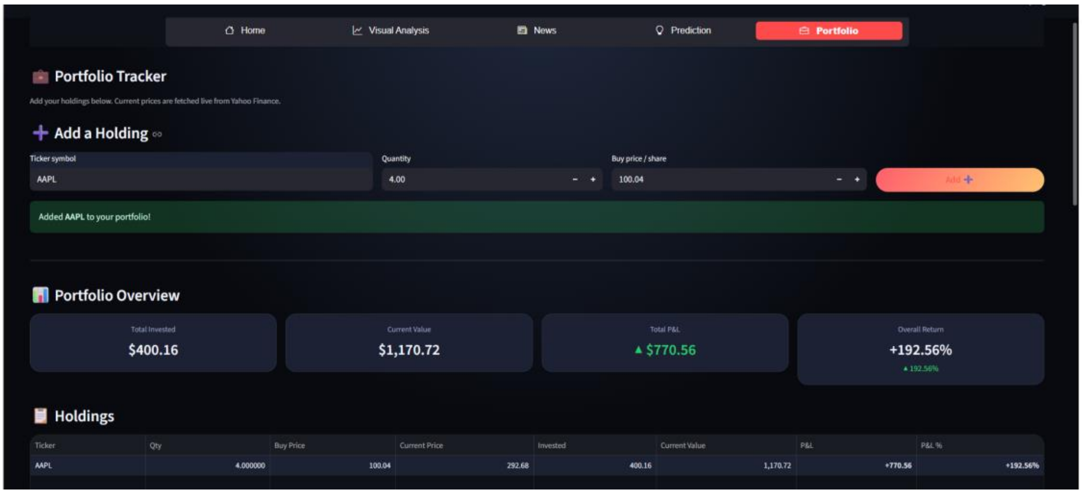

Figure 10: Iterative multi-step rollout

6.6 Sidebar

A sliding window of 18 market symbols – Gold, Silver, Oil, BTC, ETH, major indices, and individual stocks – shows the price and daily percentage change for each symbol, updated every 3 minutes along the top of the sidebar. Just below the tape, there are two ranked lists of the top 5 gainers and top 5 losers in the list of 31 US & Indian stocks for the session. There is a model controls section at the bottom that features a per-ticker retrain button that clears model files from the cache, removes the Streamlit resource cache entry and retrains the model again with epoch and loss reporting on the fly.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

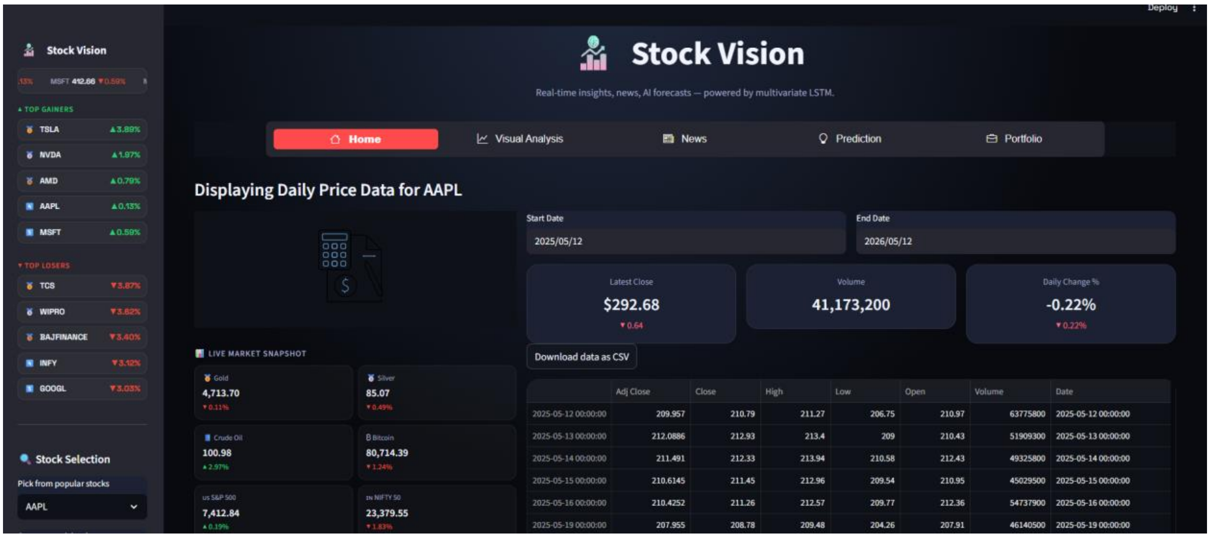

Figure 11: Iterative multi-step rollout

7. RESULTS

Several consistent conclusions emerged in the course of qualitative evaluation, conducted on a representative sample of the U.S. and Indian tickers, as detailed in Table 2 below.

Table 2: Evaluation Observations

Finding	Detail
Multivariate advantage	On trending stocks — AAPL, NVDA, RELIANCE.NS — the five-channel model anticipated volume-confirmed breakouts that the close-only baseline missed entirely
RSI contribution	When RSI exceeded 70 across consecutive window timesteps, predicted trajectories turned down ahead of actual reversals more often than the baseline
Sentiment contribution	After positive earnings-related news, retrained models produced upward-biased forecasts that matched actual opening direction more frequently
Horizon reliability	Seven-day forecasts balanced accuracy and planning utility best; thirty-day rollout accumulated error that degraded magnitude precision progressively
Indian equity support	NSE tickers trained and ran without modification; the 0.0 sentiment fallback reduced event-response sharpness, noted transparently in the UI
Portfolio computation	Weighted cost basis and live P&L computed correctly across all tested multi-purchase scenarios; 60-second cache prevented API overload

8. LIMITATIONS

8.1 Single-value sentiment

Taking the average of 50 news articles as a single number removes any time dimension from the sentiment trajectory during the session. An informative sentiment time series would be a daily sentiment time series, which would need to be a paid API tier or a custom sentiment scraping pipeline.

8.2 Rollout error accumulation

Where small errors in one prediction are multiplied over the forecast horizon by assuming each is the true forecast for the next step. This could be avoided in a direct multi-output architecture, where all the horizon values are predicted in a single forward pass, but the training objective has to differ, and more data is needed per target value.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026

DOI: 10.17148/IJARCCE.2026.15638

8.3 No fundamental data

Price discontinuities occur when a stock's price changes because of an earnings announcement, dividend adjustment or management guidance that can't be predicted from previous price and volume data. The model would have at least some warning if it included structured fundamental signals like EPS and P/B ratio.

8.4 Indian news gap

Alpha Vantage's sentiment coverage is for U.S.-listed companies. In contrast, Indian tickers are forever maintained at zero sentiment input, and thus effectively have four channels instead of five.

9. CONCLUSION

Stock Vision began with the realization that actual market players don't only look at the price chart. They aren't just watching the volume, but they are also watching the momentum reading, watching the news sentiment, and watching the trend context all at the same time, and if they aren't watching all these things, then a prediction model that sees less than what they are seeing will consistently have trouble in the conditions that matter most. The main contribution of this work is the creation of a five-channel input representation that encodes all of these dimensions, embedding the resulting model into a fully functional analytics platform, rather than as an isolated experiment. The results validate the hypothesis: including the volume, the RSI state, trend regime and news sentiment in addition to the price results in a more sensible model around breakout, reversals and news events than the price-only model. The rest of the gaps (scalar sentiment, rollout error, lack of fundamental data, lack of Indian coverage) are well understood and are therefore more of concrete steps to be taken than architectural issues.

REFERENCES

[1]. R. Chaudhary and A. Kumar, *"Advanced stock market prediction using long short-term memory networks: A comprehensive deep learning framework,"* arXiv preprint 2505.05325, 2025. https://arxiv.org/abs/2505.05325

[2]. Research Group, *"Comparative analysis of LSTM, GRU, and transformer architectures applied to equity price forecasting,"* arXiv preprint 2411.05790, 2024. https://arxiv.org/abs/2411.05790

[3]. V. Kuber, D. Yadav and A. K. Yadav, *"Univariate and multivariate LSTM applied to short-horizon stock market forecasting,"* arXiv preprint 2205.06673, National Institute of Technology Hamirpur, 2022. https://arxiv.org/abs/2205.06673

[4]. Research Group, *"Short-term equity price forecasting using a multivariate bidirectional LSTM on NIFTY 100 constituents,"*arXiv preprint 2409.14693, 2024. https://arxiv.org/abs/2409.14693

[5]. H. Lee, S. H. Lee, K. R. Lee and J. H. Kim, *"Incorporating ESG-derived sentiment and technical indicators into a deep learning framework for stock market prediction,"* Scientific Reports, Nature Portfolio, 2024. https://www.nature.com/articles/s41598-024-61106-2

[6]. Multiple Authors, *"Quantifying the contribution of market sentiment to neural network-based stock forecasting accuracy,"* Computational Economics, Springer Nature, 2024. https://link.springer.com/article/10.1007/s10614-024-10711-4

[7]. G. Chatziloizos, "Deep learning for stock market prediction using sentiment and technical analysis," *SN Computer Science*, Springer, 2024. https://link.springer.com/article/10.1007/s42979-024-02651-5

[8]. Y. Guo, "Stock price prediction based on LSTM neural network: the effectiveness of news sentiment analysis," retrieved financial news from the New York Times, combined quantitative sentiment scores with historical stock features, and confirmed that LSTM models including sentiment factors achieve higher prediction accuracy than models using only historical price data. Y. Guo, "Stock price prediction based on LSTM neural network: the effectiveness of news sentiment analysis," *IEEE Conference Publication*, IEEE Xplore, 2021. bohrium

[9]. A hybrid model integrating FinBERT-LSTM with financial news sentiment analysis was evaluated on NASDAQ-100 index stocks trained on Benzinga news articles, with results confirming that FinBERT-LSTM outperforms standalone LSTM and DNN models in prediction accuracy. Authors, "Predicting stock prices with FinBERT-LSTM: integrating news sentiment analysis," *Proceedings of the 2024 8th International Conference on Cloud and Big Data Computing*, ACM, 2024. https://dl.acm.org/doi/10.1145/3694860.3694870 Springer

[10]. R. Jawale and R. A. Deshmukh, "Forecasting future stock market values using long-term memory models and sentiment analysis," demonstrated that LSTM models integrated with sentiment analysis achieve a peak accuracy of 97% during testing, establishing LSTM and sentiment analysis as a powerful combination for accurate stock price forecasting. R. Jawale and R. A. Deshmukh, "Forecasting future stock market values using long-term memory models and sentiment analysis," in *Lecture Notes in Networks and Systems*, vol. 1389, Springer, Singapore, 2025. https://link.springer.com/chapter/10.1007/978-981-96-6066-7_1 PubMed Central

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 6, June 2026
DOI: 10.17148/IJARCCE.2026.15638

[11]. C.-R. Ko and H.-T. Chang, "LSTM-based sentiment analysis for stock price forecast," proposed using multiple factors, including news and forum article sentiment, combined with LSTM for stock price prediction, confirming that news and forum articles significantly affect price changes in the stock market. C.-R. Ko and H.-T. Chang, "LSTM-based sentiment analysis for stock price forecast," *PeerJ Computer Science*, 2021. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7959635/ Taylor & Francis Online

[12]. A novel deep learning hybrid classification model combining peephole LSTM with a temporal attention layer (TAL) was evaluated on daily data from four world indices US, UK, China, and India spanning 2005 to 2022, demonstrating superior prediction performance over CNN, LSTM, SVM, and RF benchmarks. Authors, "Enhanced prediction of stock markets using a novel deep learning model PLSTM-TAL," *PMC / National Library of Medicine*, received May 2023, accepted March 2024 https://pmc.ncbi.nlm.nih.gov/articles/PMC10963254/ arXiv