Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

High-Frequency Financial Analytics using a Decoupled Pub/Sub Architecture with Sentiment-Aware Forecasting Models

M. K. Pathak[1], Savi Dhoble[2], Janhavi Parihar[2], Aryan Gampawar[2]
[1]Department of Information Technology, All India Shri Shivaji Memorial Society's Institute of Information Technology, Pune, Maharashtra, India

Abstract - Real-time stock market analysis requires a system which can handle continuous data streams and provide accurate prediction. Our paper presents Fin-Verse, a scalable financial system, is designed to fetch live-data and give accurate predictions. The application collects live-stock data from market using Upstox API using a WebSocket based streaming. WebSocket ensures uninterrupted data flow. To improve the efficiency of our system, a dual-database strategy is used, where top 150 commonly accessed stocks are stored in primary database and other 600 plus stocks are stored in secondary database, for efficient database querying.

FinVerse improves the prediction performance by combining the numerical market data such as historical prices, sentiment analysis and trading volume. The system also includes a paper trading feature that allows user to practice stock market using fake money, helping the beginners to learn stock without any finan-cial loss. In addition to this, our system consists of an intelligent chatbot built using RAG (retrieval-augmented generation) model and the Google Gemini API, allowing users to get stock information and insights through natural language queries.

The experimental results, show that FinVerse provides a reliable analysis of data, better user-system interaction, and is also scalable. This makes it useful for modern market analysis and decision making.

Keywords: Real-Time Stock price prediction, financial forecasting, Pub/Sub ar-chitecture, LLM, Sentiment analysis, WebSocket

1 INTRODUCTION

The rapid growth of digital transaction platforms and the increasing convenience of real-time financial data have transformed the way investors analyze and participate in stock markets. New financial organizations must process large volumes of continuously changing data while providing timely insights to support **informed decision-making**. Traditional stock investigation tools often operate in inaccessible surroundings, requiring users to rely on multiple platforms for market data, predictions, and investment supervision. This disintegration leads to inadequacies and limits the effectiveness of real-time market analysis.

To address these challenges, this paper presents FinVerse, a scalable and actual financial analytics system designed to support live-stock monitoring and prediction. The proposed system employs a publish–subscribe (Pub/Sub) architecture combined with WebSocket-based streaming to ensure efficient and continuous data flow from live market sources. A dual-database strategy is adopted to optimize data retrieval by separating frequently accessed stocks from the broader market dataset. In addition, FinVerse integrates machine learning calculation techniques with sentiment analysis to improve predicting accuracy.

Beyond analysis, the system also focuses on user engagement and convenience by including a paper trading module for risk-free practice and an intelligent chatbot that provides market insights through natural language interface. By combining real-time data processing, predictive analytics, and interactive features within a unified platform, FinVerse aims to offer a complete and practical solution for current stock market analysis.

2 RELATED WORKS

The explosive increase of availability for financial data and investment environment volatility has stimulated numerous studies on intelligent stock market forecasting and decision support systems. Time-series prediction has a long history in statistics and machine learning; however classical statistical and machine learning methods were not able to capture the complicated and non-linear nature of today's financial markets. Ad-vances in deep learning, natural language processing and reinforcement learning have greatly enhanced the predictive capability of algorithms through modelling temporal dependence, investor sentiment and adaptive decision-making. Sentiment-aware and multi-modal approaches that integrate numeric market data with textual financial con-tent have been more recently investigated by scholars. Similar to this, there have been also efforts to address the practical understandability and investment results by intro-ducing real-time processing architectures and intelligent portfolio management sys-tems. However, the vast majority of prior works treat individual components inde-pendently of others and rarely take an end-to-end integrated view. This section intro-duces related work in deep learning–based forecasting, sentiment analysis, multi-modal financial intelligence and reinforcement learning to frame the contributions of the proposed FinVerse framework.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

2.1 Deep Learning–Based Financial Time-Series Forecasting

2.1.1 Standalone Machine Learning and LSTM Models
The early stock market prediction models were mainly based on traditional machine learning approaches including Artificial Neural Networks (ANN), Support Vector Re-gression (SVR) and single LSTM model. Such methodologies showed that it was pos-sible to learn from the temporal relations between historical stock prices and technical indicators. However, empirical studies showed that such models are not robust to the market volatility and they do not generalize when trading conditions change in finan-cial markets leading to large prediction errors during sharp price moves. (AI Based Stock Market Prediction)

2.1.2 Hybrid CNN–LSTM Architectures
To overcome the limitations of LSTM models, hybrid architectures referred as CNNs-LSTMs has been further proposed. CNN layers are able to capture the localized price patterns and short-term fluctuations, while LSTM layer model long-run temporal de-pendence. Research results indicate that CNN–LSTM hybrids give better predictions on both Rooted Mean Square Error (RMSE) and directional accuracy than single LSTM in the context of more volatile markets. Nevertheless, these models are still purely data-driven and do not account for external market influences. (AI Based Stock Market Prediction)

2.1.3 Generative and Transformer-Based Time-Series Models
Recent research explores generative AI models such as GANs, VAEs, and Trans-former-based architectures for financial forecasting. These models excel at learning complex data distributions and long-range dependencies, producing improved predic-tive accuracy over traditional deep learning models. Despite their performance gains, most implementations focus exclusively on price-series modeling and lack integration with sentiment or decision-making mechanisms. (Exploring Generative AI Models)

2.2 Multi-Modal Financial Intelligence Frameworks

2.3.1 Multi-Source Data Fusion for Prediction Accuracy
Multi-modal systems integrate numerical price data, technical indicators, and textual sentiment to capture a broader view of market dynamics. Research demonstrates that combining heterogeneous data sources significantly improves forecasting robustness and stability compared to single-modal approaches. Nonetheless, fusion strategies are often heuristic-based and lack adaptive learning mechanisms. (Prophetic markets Multi-modal)

2.3.2 Production-Ready Real-Time Architectures
Recent approaches focus on real-time ingesting and querying of data, employing tools such as WebSocket's, stream processing engines or low-latency pipelines. These sys-tems provide forecasts almost in real time and therefore are appropriate for intraday trading situations. Even though they are capable in engineering strength, most architec-tural designs focus on speed more than the intelligence of making decisions, and doing reasoning at the level of portfolios. (Short-term Stock Price Prediction and Stock Tech-nical Analysis System)

2.3.3 Visualization and User-Centric Decision Support
New research investigates digital-experience immersive visualization and intelligent dashboards for financial decision support, such as AR interfaces and interactive ana-lytics. Although these systems enhance interpretability and user interaction, they do not generate the desired decision strategies via learning but rather use precomputed analytics. (Immersive Financial Data Visual)

2.3 Reinforcement Learning and Portfolio Decision Systems

2.4.1 Reinforcement Learning for Trading and Portfolio Optimization
The field of RL has been used to make trading decisions automatic by discovering the optimal policy from interacting with the market. RL agents can learn to dynamically trade-off risk and return in response to market conditions. However, a wide range of RL-based systems are introduced to follow the price signals only while neglecting sen-timent-driven market psychology

2.4.2 Educational and Simulation-Based Financial Platforms
AI-based trading simulators and portfolio learning systems offer the ability to experi-ment with different investment strategies risk-free. These systems incorporate machine learning analytics and conversational agents, but focus mainly on educational purposes and do not truly connect to real-world predictive integration with sentiment and fore-casting models. (FynVerse A RAG Enhanced AI Stock)

2.4.3 Intelligent Financial Planning Systems
Innovative AI platforms for wealth management bring together predictive analytics, personalization and explainability. Although

these systems help to make user-specific investment planning, they tend not to integrate the prediction, sentiment analysis and recommendation modules together but rather optimize them separately via RL tech-niques. (InvestMate An Integrated AI-Driven)

3 PROPOSED SYSTEM

3.1 Sentiment Analysis in Financial Market Prediction

2.2.1 Traditional NLP and Machine Learning–Based Sentiment Models
Early sentiment-aware forecasting models were designed based on lexicon-based tech-niques and traditional machine learning classifiers (e.g., Random Forests, SVM) to cap-ture the sentiment of financial news headlines. These studies showed that the market sentiment has a significant influence on the direction of stock prices and reported ex-ceedingly high classification accuracy. but they do not model the contextual

significance and the long-distance dependencies in financial language. (Utilizing Sen-timent Analysis)

2.2.2 Transformer-Based Financial Language Models
Transformer-based models such as BERT and FinBERT advanced OEES by using con-textual word embeddings and attentive mechanisms. Financial-domain–informed models outperformed in identifying fine-grained changes of sentiment across earnings reports, news articles and macroeconomic commentaries. Such models enhance the sen-timent representation, but they are frequently utilized independently without being seamlessly integrated into predictive or decision-making pipelines. (Exploring Gener-ative AI Models)

2.2.3 Event-Driven and Real-Time Sentiment Integration
Advanced systems combine real-time news sentiment with market data to create event-driven predictions. Multi-modal architecture incorporating sentiment streams and price data yields significant improvements in short-term forecasting performance. However, the vast majority of studies commit to direction prediction rather than consider long-term portfolio-level results or risk conscious decision-making. (Prophetic markets Multi-modal and Short-term Stock Price Prediction)

4 METHODOLOGY

4.1 Overview and Basic understanding

The methodology of the FinVerse focuses on the implementation and design of a real-time financial system which is capable of processing heterogeneous data streams. Our Project, FinVerse, integrates real time stock market data, sentiment analysis into a sin-gle powerful dashboard referred as FinVerse.

An event driver Publisher-subscriber (Pub/Sub) model is used to ensure low latency data ingestion and processing of data. It also involves WebSocket-based continuous data streaming. The process involves 5 major stages, which are – real time data collec-tion, processing of data and standardization, extracting sentiments, preparing predictive model and creating a visual dashboard. This structured approach enables high con-sistency in monitoring market and improves the decision making.

4.2 Data Sources and Dataset used

4.2.1 Real-Time Stock Market Data

The Real-time market data is collected using Upstox API. This API provides live and historical data for National Stock Exchange (NSE) and index data.
The primary parameters include:

- Last Traded price (LTP)
- Open, High, Low, Close (OHLC) prices
- Trading Volume
- Timestamp
  The data is fetched continuously, making it efficient for real-time analysis and predic-tion.

4.2.2 Dual-Database Design for Optimization

To improve the data retrieval efficiency and performance, the system includes dual-database system strategy:

- Primary Dataset:
  Contains approximately 650 top-performing and widely used stocks. These are used frequently by the users and are used for fast querying and real time retrieval of data.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

- Secondary Dataset:
  Stores more than 64000 stocks, including all stocks which are least used or ac-cessed. This database allows larger market investigation without impacting the performance of other database
  This dual-database reduces latency, improves the scalability of the system and ensures management of real-time data.

4.2.3 Sentiment Data

Sentiment data refers to the qualitative market signals that can be mined from financial text (such as market sentiment indicators). Sentiment scores are simply changed into numerical values allowing natural joining with quantitative stock-market data.
Each sentiment score is represented on a scale:
$S \in [-1,1]$
Where:

- -1 indicates negative sentiment
- 0 indicates neutral sentiment
- +1 indicates positive sentiment

4.3 Tools, Technologies, and Execution

The system is implemented using the following tools and technologies:

Table 1. Table captions should be placed above the tables.

Component	Specification
Programming Language	Python
API Integration	Upstox API
Architecture	Publisher-subscriber
Database	Structured storage with 2 different tables
Libraries	NumPy, Pandas, ML libraries
Visualization	Graph-based dashboard
LLM Integration	Google Gemini API

4.4 Pub/Sub Architecture

The system follows a Publisher-Subscriber model to decouple data producers and con-sumers. In this architecture:

- Publishers continuously fetch live data from the Upstox API and publishes them as event.
- Subscribers include only independent module like sentiment analysis, pre-diction engine. They can see only what they subscribe to.
- The data is received asynchronously, enabling corresponding processing. To keep the data flow continuous WebSockets are used alongside Pub/Sub. Unlike tra-ditional request-response, WebSockets allow constant connections, ensuring that stock price is pushed to subscribers instantly.

Advantages of Pub/Sub + WebSockets:

- Low latency
- Scalability
- Real-time updates

This architecture improves the system's scalability and fault tolerance. The use case diagram provided below Fig (1), illustrates the interaction publishers, subscribers, and analysis modules.

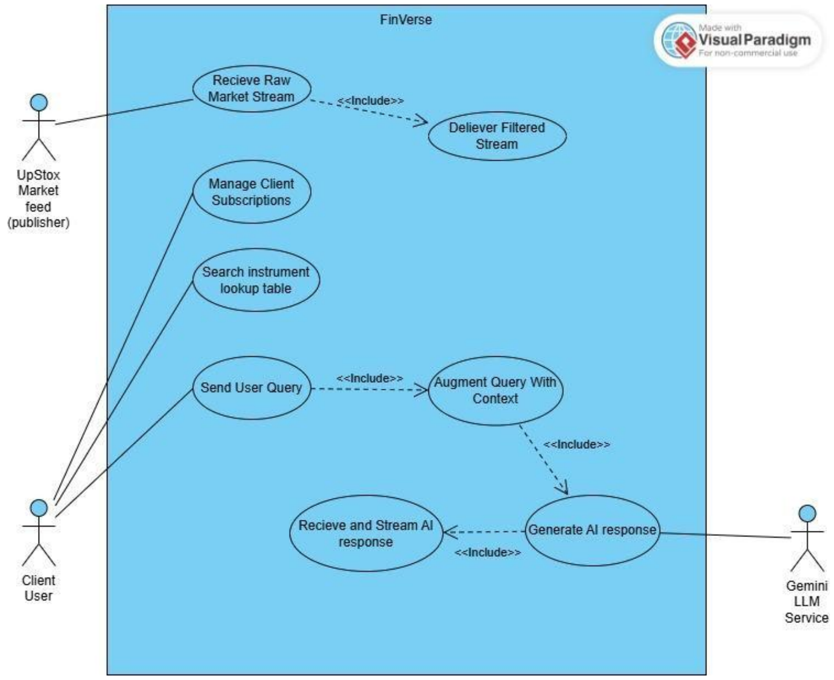

Fig. 1. Use case Diagram

4.5 Algorithms used in the system

4.5.1 Market data collection Algorithm

Algorithm 1: Real-Time Market Data Fetching

- Authenticate the API client (Upstox) using the access token provided
- Identify the instrument key for authentication
- Fetch the full market data , as per the required stock
- Parse the JSON response for any mistakes
- Store the data in appropriate database, like the primary database or Secondary database
- Then Publish the even to the subscribers. Subscribers can choose what they want to see and what they want to ignore.

4.5.2 Prediction Modelling algorithm

The prediction model combines historical data, trading volume and sentiments of the market and investors to forecast a near about price of the stocks.

$$P_{t+1} = f\,(P_t, V_t, S_t) \qquad (1)$$

Where:

- P_t= historical price
- V_t= trading volume
- S_t= sentiment score

4.5.3 Sentiment Algorithm

To enhance the prediction accuracy of the system, sentiment scores are integrated using the weighted fusion concept.

$$F_t = \alpha \cdot P_t + \beta \cdot S_t \qquad (2)$$

Where:
α and β represents the feature important weights.

Sentiments are collected from the market behavior, news, previous stock data and much more.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

4.6 Paper Trading Module

To improve the learning and engagement of users, the system includes a virtual trading feature where it provides users with fake money/ simulated money to buy and sell stocks online.

It mimics the real market behavior and is very useful for beginners to practice vari-ous trading strategies. It also reduces the financial risks as no real money is involved.

The Module allows new users to practice and learn from trying out buying or selling stocks virtually and it is solely used for educational or simulation purpose only.

4.7 Smart Chatbot using RAG model and integration of LLM (Google Gemini API)

The project features an AI-powered chatbot designed to assist users with any stock re-lated or market related queries.
A client sends a query/look-up request to the chatbot. The FastAPI Supplements the user's query with background (RAG Famework) i.e finds the best solution from a huge pile and then generates an accurate prompt to be given.

The API then generates a response using the LLM and the FastAPI streams the result back to the client

The feature enables users to obtain insights and explanations and detailed analysis of the requests, enhancing the usability and cleverness of the system.

5 RESULTS AND DISCUSSION

5.1 Experimental Setup and Execution

The proposed system, FinVerse was executed using a live stock data from the Upstox API. Using Python as a backend language, asynchronous Publish-Subscribe model was implemented, along with WebSocket's for continuous real-time data flow in our system dashboard.
The system was tested using: 1) Live NSE stock data 2) Dual database architecture
3) Sentiment Analysis and user interaction model (chatbot)

All experiments were conducted under live conditions to evaluate real-time perfor-mance, scalability and accuracy.

5.2 Graphical Analysis

5.2.1 Stock Price visualization and Prediction

Line Graphs are generated comparing the actual stock prices and predicted prices over some time.

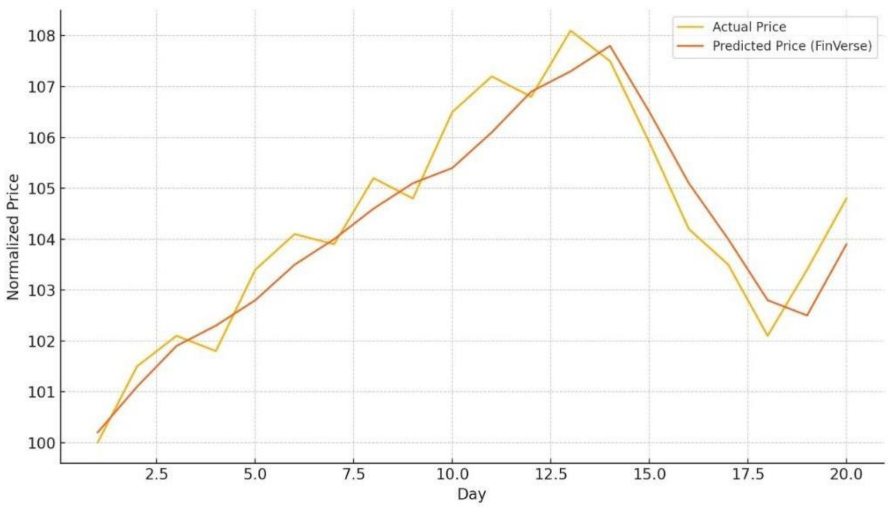

Fig. 2. Predicted price vs Actual price

The fig (2) shows the actual price vs the predicted prices using a line graph. We can say that the predicted price closely follows the actual market price, during every period of 20 days, which ensures that system captures the accurate data closely. The orange line represents the predicted price by FinVerse model.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

5.2.2 Sentiment Analysis and stock price prediction

The below map shows how likely the model predicts the correct direction, in which the price will move (UP or DOWN). This helps to calculate the precision of our system:

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

The Precision is how many time the UP value is actually predicted correctly. Integrating sentiment Features reduced the false positives during frequent changing periods, im-proving the overall prediction and reliability of the system.

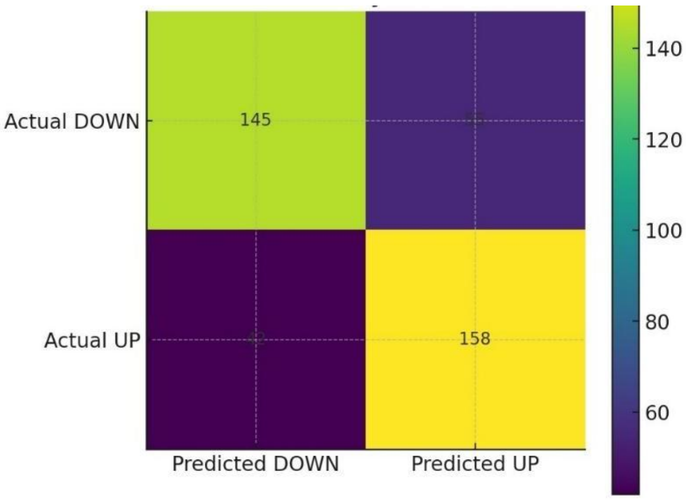

Fig. 3. Confusion Matrix

The above Heatmap Fig (3) is a confusion matrix, which helps to calculate the accu-racy and precision of the system. The vertical direction shows the actual UP/DOWN direction and the horizontal line shows the predicted UP/Down direction.

5.3 Testing and Evaluation Metrics

Test Case	**Input**	**Expected Output**	**Status**
API Data Fetch	Instrument Key	Market Data Retrieved	Success
Database Directing	Check Stock Type	Correct DB Access given	Success
Prediction Output	Market Data collected	Forecast Generated accurate result	Success
Virtual Trading	Fake/ Paper money	Trade Executed for educational purposes	Success
Chatbot Querying /lookup request	User Prompt is given	Relevant Response given by RAG and LLM (Google Gem-ini model)	Success

5.4 Final Outputs and discussion

The live data was continuously captured using Upstox API using WebSocket based streaming, which ensured that the data flow is uninterrupted. The execution of our

system successfully ensures the integration of real-time data, predictive analysis and visualization of our financial dashboard.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

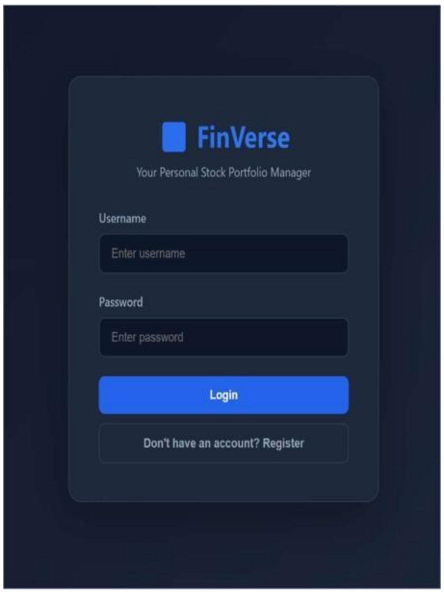

Fig. (4a) Login form

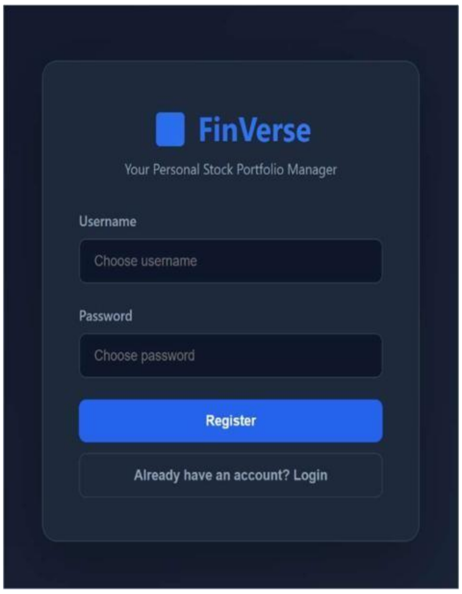

Fig. (4b) Registration form

Figure 4a and 4b represents the login and registration interface, ensuring secure authen-tication and system access. This module ensures that only registered users can access the features like stock prediction, portfolio optimization and Gemini based chatbot.

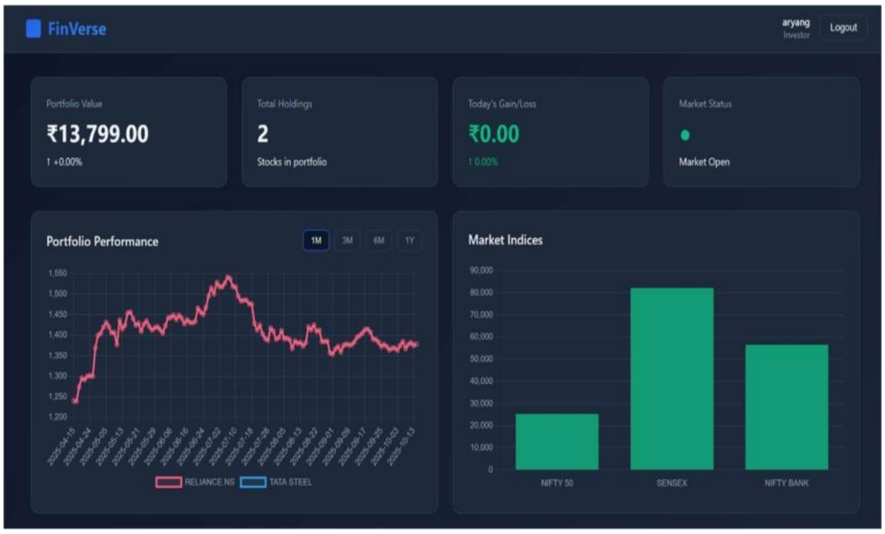

Fig. (4c) Dashboard displaying market status

Fig. 4c shows the FinVerse Dashboard displaying the live market data, where stock prices are predicted based on historical data and sentiment analysis. The continuous up-dates are observed in the dashboard because of continuous streaming using WebSocket using the Upstox API. The Pub/Sub model helps to effectively show the events to the subscribers.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

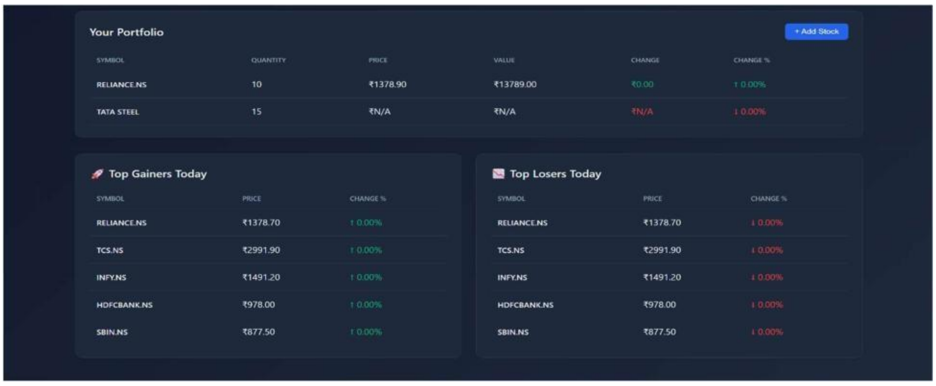

Fig. (4d) Stock insights

The fig. 4d shows the stock insights module, which provides a view of all the selected stocks, including historical trends, predicted price and sentiments. The percentage change closely predicts the real market movements.

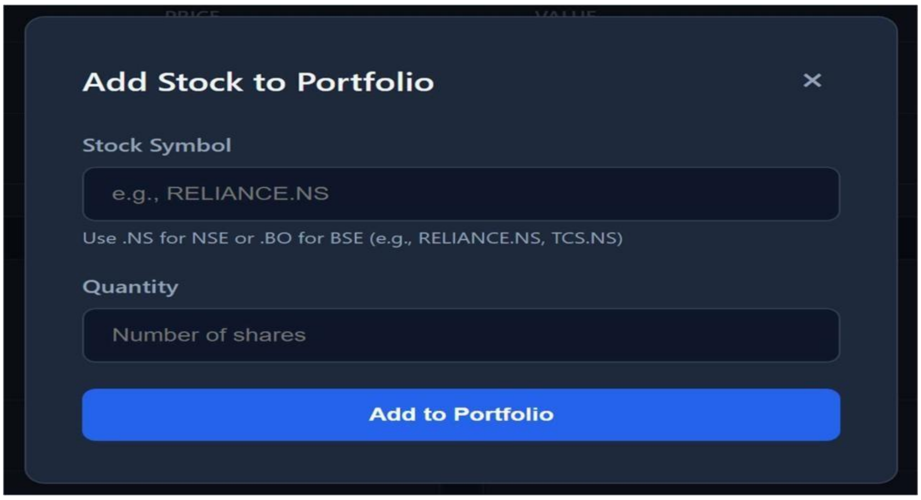

Fig. (4d) Adding stocks

Fig 4d shows the add new stock feature, where users and dynamically add stocks to the system for tracking. The stock symbol is used to add stock like RELIANCE.NS as

shown in image and the Quantity defines how much stocks to add to your portfolio. This helps users to track their portfolio, and also add new stocks they are interested in. Users have complete control over what they want to track.

6 CONCLUSION

This paper presented FinVerse, a real-time financial analytics system that integrates live stock market data, predictive modeling, and interactive user features. The use of WebSocket-based streaming and a publish–subscribe architecture enables efficient, low-latency data processing, while a dual-database approach improves scalability and data retrieval performance. The multimodal prediction model combining historical data and sentiment analysis achieved reliable forecasting accuracy. Features such as paper trading and an LLM-based chatbot further enhance user engagement and acces-sibility. Overall, FinVerse demonstrates an effective and scalable solution for real-time stock market analysis and decision support.

REFERENCES

[1] T. Ding, S. Qian, J. Cao, G. Xue and M. Li, "SCSL: Optimizing Matching Algorithms to Improve Real-time for Content-based Pub/Sub Systems," 2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS), New Orleans, LA, USA, 2020, pp. 148-157, doi: 10.1109/IPDPS47924.2020.00025

[2] A. K. Singh, A. Singh, A. Ishan, A. Kumar, G. K. Dhari and B. Kumar, "AI Based Stock Market Prediction," 2024 1st International Conference on Advances in Computing, Com-munication and Networking (ICAC2N), Greater Noida, India, 2024, pp. 769-771, doi: 10.1109/ICAC2N63387.2024.10895866.

[3] S. Meganathan, "Immersive Financial Data Visualization Using Augmented Reality Smart Glasses: A Framework for Real-Time Investment Decision Support," 2025 2nd International Conference on Electronic Circuits and Signaling Technologies (ICECST), Petaling Jaya, Malaysia, 2025, pp. 940-946, doi: 10.1109/ICECST66106.2025.11307512.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

[4] W. Ding, H. Wang, N. Peng, Y. Xiao and Z. Liu, "Stock Technical Analysis System Based on Real-Time Stream Processing," 2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2017, pp. 373-377, doi: 10.1109/ISCID.2017.218.
[5] S. Banoth and S. Yadala, "A Hybrid Deep Learning Framework for Stock Price Forecasting with Sentimental Analysis," 2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, India, 2024, pp. 467-471, doi: 10.1109/SMART63812.2024.10882215.
[6] Talamala, U. R, F. Praveena, V. Manoranjithem, S. Mastan and A. Perada, "Enhancing Stock Market Predictions through CNN-Based Hybrid Deep Learning Models," 2025 5th Asian Conference on Innovation in Technology (ASIANCON), PIMPRI, India, 2025, pp. 1-6, doi: 10.1109/ASIANCON66527.2025.11280763.