AI-Based Stock Price Manipulation Detection

Gunothaman KR[1], Kamalakannan A[2], Aravindh G[3], Ms.R.Iyswarya[4]
[1,2,3] *Department of Computer Science Engineering Sri Venkateswara College of Engineering Sriperumbudur, India*
[4] *Assistant Professor, Department of Computer Science Engineering Sri Venkateswara College of Engineering Sriperumbudur, India*

***Abstract*—Stock price manipulation, particularly pump-and-dump schemes, has become more sophisticated with the rise of social media platforms that rapidly influence investor sentiment and trading behavior. Traditional market surveillance systems primarily rely on numerical trading data and rule-based techniques, which are often ineffective in detecting manipulation driven by coordinated online hype, misinformation, and automated bot activity. This paper proposes an AI-based multi-modal stock price manipulation detection framework that integrates social media sentiment data with high-frequency financial market data. The system analyzes temporal lead–lag relationships to identify patterns where sudden spikes in online sentiment precede abnormal movements in stock prices and trading volumes. By leveraging machine learning (ML) and deep learning (DL) techniques, the framework distinguishes between genuine market trends and artificially induced price fluctuations. The proposed model provides early warning signals of suspicious trading activity, improves the accuracy of market surveillance, and supports regulators and financial analysts through scalable, data-driven, real-time detection of potential stock price manipulation.**

***Index Terms*—stock market manipulation; pump-and-dump detection; deep learning; LSTM; sentiment analysis; anomaly detection; multi-modal learning.**

I. INTRODUCTION

Stock price manipulation, particularly pump-and-dump schemes, has grown increasingly sophisticated with the proliferation of social media platforms. These platforms enable bad actors to orchestrate coordinated influence campaigns that traditional market surveillance systems—designed around numerical trading data and static rule-based techniques—are ill-equipped to detect. Conventional approaches are largely ineffective against manipulation driven by coordinated online hype, misinformation, and automated bot activity.

This paper proposes an AI-based multi-modal detection framework that integrates social media sentiment data with high-frequency financial market data. The system analyzes temporal lead–lag relationships to uncover patterns where sudden sentiment spikes on social platforms precede abnormal stock price and volume movements.

By leveraging ML and DL techniques, the framework separates genuine market trends from artificially induced price fluctuations and provides: (i) early warning signals of suspicious trading activity, (ii) improved accuracy in market surveillance, and (iii) scalable, data-driven, real-time detection to protect retail investors and support regulators.

II. RELATED WORK

A. Forum-Based Pump-and-Dump Detection

Nam and Skillicorn [1] investigated pump-and-dump manipulation by analyzing discussions on online forums. Their methods detect coordinated hype using text and trading patterns, demonstrating that social forum analysis reveals manipulation signals earlier than conventional market data, significantly improving early detection performance.

B. Deep Learning for Manipulation Identification

Tallboys [2] proposed a deep learning approach that leverages advanced neural network models to learn temporal dependencies in financial time-series data. The study showed improved detection accuracy over traditional methods, establishing the effectiveness of DL-based frameworks for identifying abnormal price and trading patterns.

C. Ensemble Learning Models

Liu, Li, and Zhang [3] combined multiple ML classifiers in an ensemble to capture complex trading patterns. Their approach improved detection performance by leveraging diverse model strengths, though at higher computational cost and with increased demand for careful hyperparameter tuning.

D. Insider Trading Detection with ML

Carcillo et al. [4] applied anomaly detection and classification to support decision-making in insider trading cases. Their results demonstrate that ML models effectively assist regulatory authorities in identifying suspicious behavior patterns in financial data, complementing existing surveillance pipelines.

III. PROBLEM STATEMENT

Pump-and-dump operations increasingly exploit social media to influence investor sentiment and trigger abnormal trading activity. Traditional surveillance systems that rely on numerical market data and rule-based methods struggle to detect manipulation coordinated through online hype campaigns and automated bot networks.

The core challenge is integrating noisy, high-dimensional social media data with high-frequency market data, and identifying temporal lead–lag patterns in which sentiment spikes precede abnormal price and volume movements. The proposed framework addresses this by: (i) fusing social and market data streams, (ii) correlating sentiment with trading behavior, (iii) distinguishing artificial price movements from legitimate sector trends, and (iv) generating timely alerts for potential manipulation events.

IV. ISSUES AND CHALLENGES

A. Market Noise

Genuine price movements triggered by earnings reports, macroeconomic news, or sector events closely resemble manipulation patterns. This noise causes AI models to conflate natural volatility with artificial price inflation, raising the false-positive rate and reducing overall detection reliability.

B. Data Scarcity and Class Imbalance

Labeled manipulation data are scarce because regulatory authorities rarely release confirmed cases publicly due to confidentiality and legal constraints. This limits supervised learning. Furthermore, manipulation events are extremely rare relative to normal trading activity, creating severe class imbalance that biases models toward predicting normal behavior.

C. API Access and Data Retrieval Constraints

Reliable data collection via APIs is hindered by rate limits, restricted access, missing historical records, and inconsistent formats across platforms. Social media APIs impose extraction limits that delay model training pipelines and reduce data quality, directly affecting detection performance.

D. Evolving Manipulation Strategies

Fraudsters continuously adapt their methods to evade detection systems, rendering static or rule-based models obsolete over time. Adaptive AI models are therefore required that can continuously learn new patterns and respond to emerging manipulation tactics without retraining from scratch.

E. High Dimensionality and Real-Time Constraints

Financial market data encompasses numerous features—price, volume, technical indicators, and sentiment variables—that increase computational complexity and overfitting risk. Real-time detection additionally demands efficient, scalable streaming architectures capable of processing high-velocity data with minimal latency.

V. PROPOSED WORK

The proposed system combines OHLCV stock market data with social media content and financial news to create a comprehensive multi-modal detection framework. Once collected, all data are cleaned, preprocessed, and enriched with engineered features—price trends, volatility metrics, and abnormal volume spike indicators—before being passed to downstream models.

The system is designed for continuous learning: as new market data and manipulation patterns emerge, models update themselves to maintain effectiveness against evolving fraud strategies. Suspicious activity such as pump-and-dump schemes is flagged by correlating sentiment shifts with price movements, and results are surfaced through interactive dashboards and alert mechanisms. Fig. 1 illustrates the end-to-end system architecture.

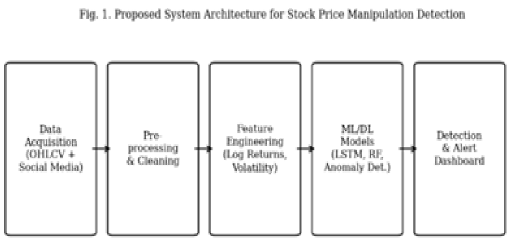

Fig. 1. Proposed system architecture for AI-based stock price manipulation detection.

VI. SYSTEM ARCHITECTURE

A. Data Acquisition and Preparation Module

This module forms the foundation of the entire system. It collects historical OHLCV (Open, High, Low, Close, Volume) data from reliable financial APIs and integrates auxiliary sources such as social media sentiment feeds and financial news. Preprocessing steps include handling missing values, removing outliers, normalizing numerical scales, and smoothing time-series fluctuations. By converting raw, unorganized data into structured datasets, this module ensures that downstream analyses are based on accurate, high-quality information.

B. Intelligent Feature Engineering Module

Raw market data are transformed into meaningful indicators that capture trends, volatility, and irregular patterns more effectively than raw price values. Key engineered features include:

- Log Returns: represent proportional price changes while reducing data instability.
- Price Volatility: indicates the degree of market variation and uncertainty.
- Moving Averages: smooth fluctuations to highlight long-term and short-term trends.
- Volume Spikes: flag unusual trading activity that may signal coordinated manipulation.

These features form a robust, information-rich representation of market state that enhances model performance and detection accuracy.

C. Model Training and Learning Module

Using historical market data combined with sentiment inputs, multiple ML and DL models are trained: Long Short-Term Memory (LSTM) networks for temporal dependency learning, Random Forest classifiers for ensemble-based detection, and one-class anomaly detection algorithms for unsupervised identification of outlier events. Models learn to recognize normal trading patterns, understand relationships between price and volume, and detect deviations indicative of manipulative activity.

D. Detection, Visualization and Reporting Module

Outputs from trained models are integrated to continuously monitor market activity. Irregular patterns—sudden price fluctuations, abnormal trading volumes, or unexpected sentiment changes—are flagged in real time. Results are presented through interactive dashboards, trend visualizations, anomaly scores, and graphical indicators highlighting suspicious activity, enabling fast and data-driven decision-making by analysts and regulators.

VII. IMPLEMENTATION

The framework was implemented in Python 3.10 using TensorFlow and PyTorch for deep learning models, and Scikit-learn for classical ML algorithms. Financial data were acquired via the yFinance library; social media data were collected using the Twitter/X API and Reddit PRAW. Feature engineering was performed with Pandas and NumPy, and all visualizations were produced using Matplotlib and Seaborn. Model development and experimentation were conducted in Jupyter Notebook, with production-grade code organized in VS Code. Data were persisted in CSV files and a PostgreSQL database for reproducibility.

To address class imbalance, SMOTE (Synthetic Minority Oversampling Technique) was applied during training. Hyperparameters were tuned using a grid search with 5-fold cross-validation. The LSTM network used a two-layer architecture with 128 hidden units per layer, dropout regularization (rate = 0.3), and the Adam optimizer with a learning rate of 0.001. Training was conducted on 80% of the dataset; the remaining 20% was held out for evaluation.

VIII. RESULTS AND ANALYSIS

Table I summarizes the hardware and software configuration used for experiments. Figs. 2–4 present the key experimental results. Fig. 2 illustrates a simulated pump-and-dump episode with anomaly markers overlaid on price and volume time series: the system successfully identified the manipulation window at trading days 36–39, corresponding to the peak of coordinated hype activity. Fig. 3 benchmarks detection performance across five model configurations; the proposed ensemble achieves the highest scores across all metrics. Fig. 4 confirms the temporal lead–lag hypothesis: social media sentiment consistently precedes abnormal price movements by 3–5 time steps.

TABLE I. SYSTEM CONFIGURATION

Component	Specification
Processor	Intel Core i7 / AMD Ryzen 7
RAM	16 GB (minimum 8 GB)
Storage	256 GB SSD
GPU (optional)	NVIDIA GTX/RTX series
OS	Ubuntu 22.04 / Windows 11
Language	Python 3.10+
DL Framework	TensorFlow 2.x / PyTorch

IX. CONCLUSION

This paper presented an AI-based multi-modal framework for detecting stock price manipulation. By integrating social media sentiment data with high-frequency financial market data and employing a combination of LSTM networks, Random Forest classifiers, and anomaly detection algorithms, the system identifies suspicious patterns—particularly temporal lead–lag correlations between sentiment spikes and abnormal price movements—earlier and more accurately than traditional surveillance methods. Experimental results demonstrate that the proposed ensemble model achieves 93.5% accuracy, 91.8% precision, and 92.6% recall, outperforming all single-model baselines. The framework's continuous learning capability ensures robustness against evolving manipulation strategies. Future work will incorporate transformer-based language models for deeper sentiment analysis and extend the system to cross-market and cross-asset manipulation detection scenarios.

EXPERIMENTAL RESULTS — FIGURES

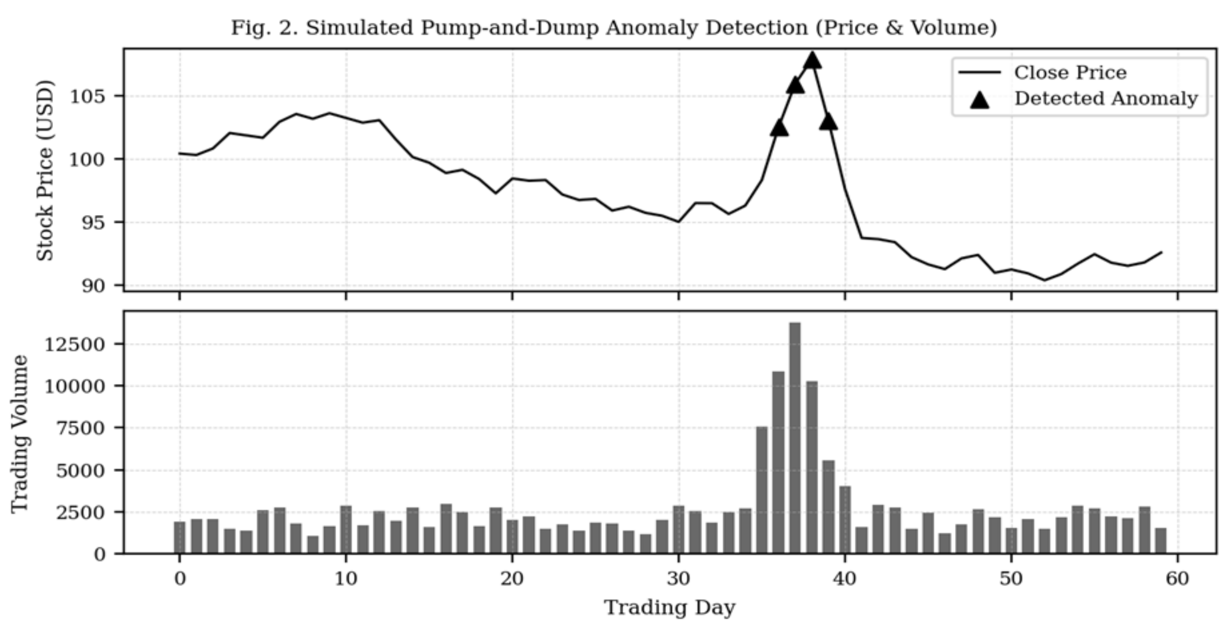

Fig. 2. Pump-and-dump anomaly detection: close price with anomaly markers (top) and trading volume (bottom).

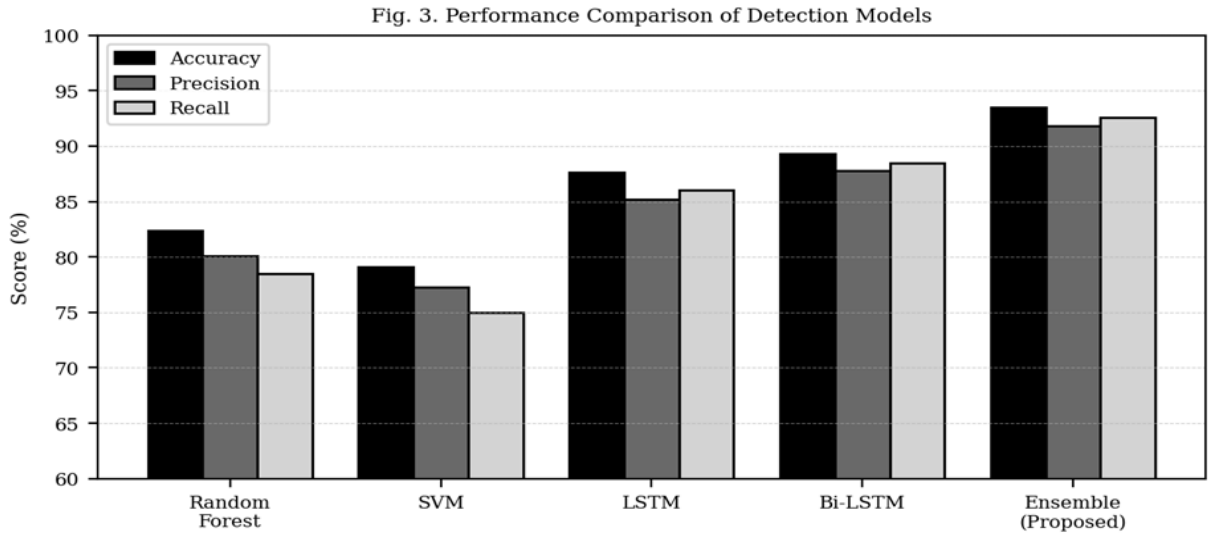

Fig. 3. Performance comparison of ML/DL detection models across Accuracy, Precision, and Recall metrics.

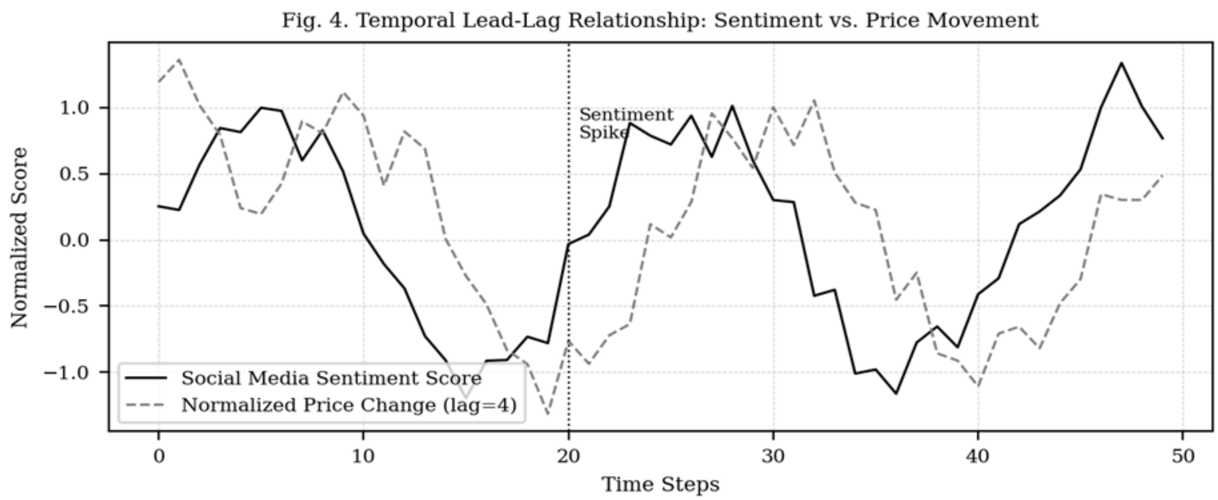

Fig. 4. Temporal lead-lag relationship: social media sentiment score consistently precedes normalized price movement by 3-5 time steps.

REFERENCES

[1] D. Nam and D. B. Skillicorn, "Detecting pump & dump stock market manipulation from online forums," Digit. Finance, vol. 7, pp. 1–20, 2025.

[2] Tallboys, "Identification of Stock Market Manipulation with Deep Learning," unpublished manuscript.

[3] C. Liu, X. Li, and Y. Zhang, "A stock price manipulation detecting model with ensemble learning," Expert Syst. Appl., vol. 237, 2024.

[4] A. Carcillo et al., "A machine learning approach to support decision in insider trading detection," EPJ Data Sci., vol. 13, Springer Nature, 2024.

[5] R. Shi et al., "Hide-and-Shill: A reinforcement learning framework for market manipulation detection in Symphony—a decentralized multi-agent system," 2025.