5. Nvidia-Powered Supercomputer Transforms Weather Research / University of Florida. – Gainesville, FL : University of Florida Press, 2023. – P. 6–9.

UDC 336.76:004.85

RESEARCH ON STOCK PRICE PREDICTION BASED ON LSTM-TRANSFORMER FUSION MODEL IN FINANCIAL MANAGEMENT

Ailila·Kadierbieke, master's degree student,
Sednina M. A., senior lecturer
Belarusian National Technical University
Minsk, Belarus

Abstract. Stock price prediction is a challenging task due to its non-linearity, volatility and temporal dependence that traditional models are hard to fully capture. This study proposes an LSTM-Transformer fusion model, combining LSTM's strength in long-term temporal feature extraction and Transformer's advantage in global correlation capture. Comparative experiments on multi-source financial data show the model outperforms single deep learning and traditional models, providing reliable support for investment decision-making.

Keywords: Stock Price Prediction, LSTM-Transformer, Fusion Model; Deep Learning, Temporal Feature, Financial Data.

1. Introduction

Against the backdrop of accelerated global financial market integration, stock prices operate as a complex dynamic system, shaped by the interplay of micro-level transaction dynamics (e. g., daily trading volume and price fluctuations), investor sentiment shifts (e. g., social media optimism/pessimism), macroeconomic policy adjustments (e. g., interest rate changes), and cross-border market spillover effects (e. g., U. S. stock market volatility impacting A-shares). For investors, accurate short-term stock price prediction (e. g., forecasting the next day's closing price) is critical for optimizing portfolio allocation and mitigating downside risks; for regulatory authorities, it serves as a key tool for identifying abnormal market activities and maintaining financial stability. However, the inherent strong nonlinearity of stock prices – exacerbated by sudden events like geopolitical conflicts or policy announcements – and complex temporal dependencies between historical and current data have long posed challenges to traditional prediction models. These models often fail to capture the full complexity of market dynamics, creating an urgent demand for adaptive analytical frameworks that integrate both sequential feature extraction and global correlation capture.

1.1. Research Background.

In recent years, the global financial landscape has witnessed heightened volatility, driven by factors such as post-pandemic economic recovery disparities, monetary policy shifts (e. g., the U. S. Federal Reserve's 2023–2024 interest rate hikes), and the rapid expansion of algorithmic trading (accounting for over 50 % of A-share market transactions in 2024). In this context, stock prices no longer follow simple linear trends but exhibit irregular fluctuations influenced by two types of data: (1) Quantifiable data: Historical transaction data (e. g., daily opening price, trading volume), quarterly financial reports, and macroeconomic indicators (e. g., GDP growth rate, CPI); (2) Unstructured data: Investor sentiment from social platforms (e. g., Eastmoney stock bars, Weibo), and financial news headlines (e. g., policy announcements on industry regulation).

For individual investors, traditional technical analysis tools (e. g., moving averages, relative strength indices) cannot process high-dimensional multi-source data in real time, leading to suboptimal decision-making. For institutional investors (e. g., hedge funds), even a 1% error in next-day closing price prediction can result in millions of yuan in losses. Notably, the interconnectedness of global markets means that a single event – such as a supply chain disruption in Southeast Asia or a European Central Bank policy announcement – can trigger ripple effects across stock exchanges worldwide. This interdependence further increases prediction complexity, as models must now account for cross-market correlations in addition to internal market dynamics.

Traditional statistical models (e. g., ARIMA) rely on assumptions of stationarity and linearity, making them ineffective at capturing sentiment-driven price jumps; traditional machine learning models (e. g., SVM) struggle to integrate macroeconomic data with transaction data; and single deep learning models (e. g., standalone LSTM) often overlook global cross-stock correlations. Thus, developing a framework that can: (1) handle nonlinearity and long-term temporal dependencies; (2) fuse multi-modal data (transaction, sentiment, macroeconomic); and (3) capture global feature correlations has become a critical priority in financial research.

1.2. Research Status.

The field of stock price prediction has evolved through three methodological phases, with distinct limitations that highlight clear research gaps:

1.2.1. Traditional Statistical Methods.

This phase centered on linear time series models, with the Autoregressive Integrated Moving Average (ARIMA) and its variants (e. g., SARIMA) as the most representative. ARIMA predicts future prices by fitting autoregressive (AR) and moving average (MA) components to stationary historical data, performing well in stable market conditions (e. g., low-volatility periods for blue-chip stocks). However, its performance degrades significantly in non-stationary and nonlinear environments – common in modern markets – since it cannot account for sudden price volatility (e. g., a 5 % daily drop driven by negative news). Even extended models like GARCH (Generalized Autoregressive Conditional Heteroskedasticity), which address volatility clustering, still rely on linear assumptions and fail to model deep correlations between macroeconomic indicators (e. g., CPI) and stock prices.

1.2.2. Traditional Machine Learning Methods.

This phase introduced nonlinear algorithms such as Support Vector Machines (SVM), Random Forest, and Gradient Boosting Machines (GBM). SVM uses kernel functions to map data into high-dimensional spaces, enabling it to model nonlinear relationships between trading volume and stock prices; Random Forest reduces overfitting by combining multiple decision trees, improving robustness for mid-cap stock prediction. However, these methods have two critical flaws: (1) They treat sequential data as independent, making it impossible to capture long-term temporal dependencies (e. g., how 30-day price trends impact future prices); (2) They require manual feature engineering, which relies heavily on domain expertise and often omits hidden features (e. g., sentiment-driven price momentum).

Studies (e. g., Li et al., 2023) show that traditional machine learning models using only transaction data have a Mean Absolute Error (MAE) 18–22 % higher than models integrating multi-modal data, highlighting their data limitation.

1.2.3. Single Deep Learning Methods.

Advancements in deep learning led to models tailored for time series: (1) LSTM: Addresses the vanishing gradient problem of RNNs using input/forget/output gates, enabling it to capture long-term temporal dependencies (e. g., predicting next-day prices using 20-day historical data). However, its sequential processing nature struggles to capture global correlations (e. g., how a tech stock's price relates to the CSI 300 Index); (2) Transformer: Uses self-attention mechanisms for parallel data processing, excelling at global feature capture (e.g., cross-industry price linkages). Yet, it lacks LSTM's fine-grained temporal memory, leading to higher MAE (15–17 % higher, per Wang et al., 2024) in short-term (1–3 day) prediction compared to LSTM.

Existing fusion models (e. g., Zhang et al., 2023) mostly adopt parallel fusion (LSTM and Transformer process data simultaneously), which fails to leverage LSTM's strength in preprocessing sequential features before global correlation extraction. Additionally, over 60 % of deep learning studies (per a 2024 survey in *Journal of Financial Data Science*) rely solely on transaction data, ignoring the complementary value of sentiment and macroeconomic data.

Key Research Gaps: (1) Lack of sequential fusion models that first use LSTM to extract temporal features, then Transformer to capture global correlations; (2) Insufficient quantification of how multi-modal data (transaction + sentiment + macroeconomic) improves prediction accuracy compared to single-modal data; (3) No systematic validation of model performance across stocks of different market capitalizations (large/medium/small-cap) and industries.

1.3. Research Questions and Objectives.

1.3.1. Research Questions.

Building on the identified gaps, this study focuses on three actionable, scenario-specific research questions:

1. Can a sequential LSTM-Transformer fusion model (LSTM extracts temporal features first, then Transformer captures global correlations) outperform traditional models (ARIMA, SVM), single deep learning models (standalone LSTM, standalone Transformer), and parallel fusion models in next-day closing price prediction of A-share stocks?

2. How do different multi-modal data combinations – (a) transaction data only, (b) transaction + sentiment data, (c) transaction + sentiment +

macroeconomic data – impact the fusion model's performance? Specifically, does adding sentiment (from Eastmoney bars) and macroeconomic data (GDP, CPI) reduce MAE and RMSE compared to single-modal data?

3. Does the fusion model maintain high prediction accuracy across A-share stocks of different market capitalizations (large-cap: e. g., ICBC; medium-cap: e. g., BOE Technology; small-cap: e. g., Jiangsu Leyou) and industries (finance, technology, consumer)?

1.3.2. Research Objectives.

To answer these questions, the study sets three reproducible, deliverable-focused objectives:

1. Construct a sequential LSTM-Transformer fusion model: Design an architecture where LSTM layer (128 hidden units, dropout rate 0.2) processes 20-day historical data to extract temporal features, and Transformer encoder (4 attention heads, feed-forward dimension 512) processes LSTM outputs to capture global correlations. Deliverables: Detailed architectural diagram, open-source TensorFlow code, and parameter tuning report.

2. Validate model superiority via controlled experiments: Data includes 10 representative A-shares (3 large-cap, 4 medium-cap, 3 small-cap) from 3 industries, with data spanning 2019–2024 (1 200+ trading days); baselines cover ARIMA, SVM, Random Forest, standalone LSTM, standalone Transformer, parallel LSTM-Transformer; metrics include MAE, MSE, RMSE, MAPE (Mean Absolute Percentage Error). Deliverable: Statistical report with t-tests ($p<0,05$) to validate significant accuracy improvements.

3. Quantify multi-modal data value: Conduct three experiments using the 10 stocks' data, comparing model performance across the three data combinations (a/b/c). Deliverable: Analysis report identifying the optimal data combination and calculating the marginal MAE reduction from adding sentiment (≈ 8–10 %) and macroeconomic data (≈ 5–7 %).

1.4. Significance of the Study.

1.4.1. Theoretical Significance.

1. Innovate fusion model architecture for financial time series: This study proposes a sequential fusion logic (LSTM → Transformer) that addresses the limitations of parallel fusion models, providing a new template for balancing temporal dependency capture and global correlation extraction in stock price prediction.

2. Deepen multi-modal data fusion theory in finance: By quantifying the marginal value of sentiment and macroeconomic data, the study clarifies how unstructured sentiment data (reflecting market psychology) and

structured macroeconomic data (reflecting systematic risks) complement transaction data – filling the gap in existing research that lacks quantitative analysis of multi-modal value.

3. Supplement cross-scale/industry model generalization theory: Validating performance across large/medium/small-cap stocks and industries extends the understanding of deep learning model adaptability in heterogeneous financial markets, addressing the lack of generalization studies in current literature.

1.4.2. Practical Significance.

1. For investors: The model provides accurate next-day closing price predictions and explains why prices fluctuate (e. g., "price rise driven by positive sentiment + GDP growth"), enabling individual investors to adjust holdings (e. g., buying undervalued stocks predicted to rise) and institutional investors to optimize algorithmic trading strategies.

2. For regulatory authorities: The model serves as an early warning tool: deviations between predicted and actual prices (e. g., a 7 % jump not justified by multi-modal data) signal potential market manipulation. Additionally, it helps regulators assess policy impacts (e. g., how interest rate hikes affect small-cap stocks) to formulate targeted supervision.

3. For financial institutions: Hedge funds can integrate the model into high-frequency trading systems to reduce transaction costs; risk management departments can use its predictions to calculate Value at Risk (VaR) more accurately, enhancing resilience during market downturns.

The model's framework is also adaptable to other financial time series tasks (e. g., commodity price forecasting, foreign exchange rate prediction), extending its practical value beyond stock markets.

1.5. Structure of the Paper.

To help readers quickly grasp the overall logic of the study and clarify the connection between each chapter and the research objectives, this paper is organized into six chapters, forming a closed loop of "theoretical foundation → technical implementation → result verification → discussion and summary":

Chapter 1: Introduction: This chapter clarifies the research background and practical demand of stock price prediction, sorts out the limitations of existing research methods and core research gaps, puts forward specific research questions and objectives, expounds the theoretical and practical significance of the study, and finally outlines the overall structure of the paper to lay the foundation for subsequent content.

Chapter 2: Literature Review: This chapter systematically reviews the research progress of traditional statistical methods, traditional machine learning methods, and single deep learning methods in stock price prediction, focuses on the application status and controversies of LSTM-Transformer fusion models, and summarizes the consensus and unresolved gaps in existing research, providing theoretical support for the design of the study's model and data strategy.

Chapter 3: Data and Methodologies: As the technical core of the study, this chapter details the selection, source, and preprocessing process of research data (covering 10 A-share stocks), designs multi-modal feature indicators (transaction, sentiment, macroeconomic), constructs the sequential LSTM-Transformer fusion model and comparative baseline models, and specifies the model training, optimization strategies, and evaluation metrics to ensure the reproducibility of experiments.

Chapter 4: Results and Analysis: This chapter presents experimental results, including descriptive statistics of research data, performance comparison between the fusion model and baseline models, analysis of the impact of different multi-modal data combinations on model performance, and verification of the model's generalization ability across stocks of different market capitalizations and industries, so as to answer the proposed research questions.

Chapter 5: Discussion: This chapter interprets the experimental results in depth, analyzes the internal mechanism of the fusion model's superior performance and the value of multi-modal data fusion, objectively points out the limitations of the study (such as sample size and prediction cycle), and puts forward feasible directions for future research.

Chapter 6: Conclusion: This chapter summarizes the core conclusions of the study, reaffirms the innovative contributions of the sequential LSTM-Transformer fusion model and multi-modal data strategy, and clarifies the theoretical and practical implications of the research, forming a complete summary of the entire study.

Readers can locate key content according to their needs: for technical details, focus on Chapter 3; for experimental conclusions, focus on Chapter 4; for academic discussions, refer to Chapter 5.

1. Literature Review.

1.1. Traditional Stock Price Prediction Methods.

1.2. Deep Learning-Based Stock Price Prediction Methods.

1.3. Multimodal Data Fusion in Stock Price Prediction.