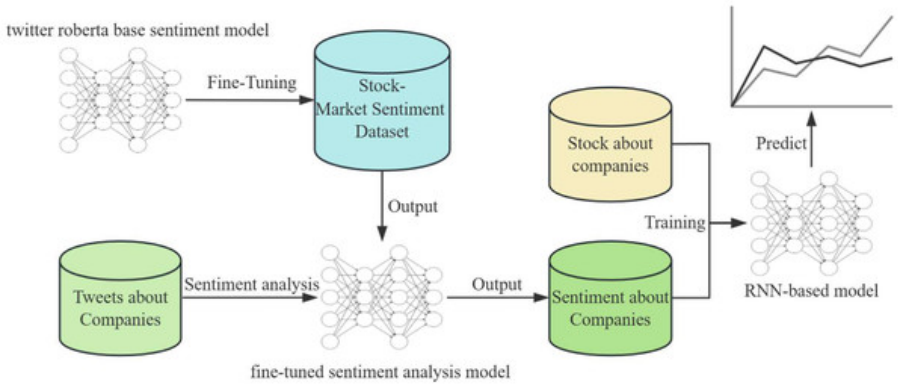
Deep learning in finance assessing twitter sentiment impact and prediction on stocks

Raj Tripathi (16014223062) — Batch A2
K. J. Somaiya School of Engineering, Vidyavihar

Introduction

The research paper, "Deep Learning in Finance: Assessing Twitter Sentiment Impact and Prediction on Stocks" by Kaifeng Guo and Haoling Xie, investigates the relationship between public sentiment on Twitter and stock market fluctuations and proposes a deep learning framework to predict stock prices by integrating sentiment information [cite: 15, 16, 20, 97]. The study highlights how social media sentiment reflects investor behavior and demonstrates that incorporating sentiment into predictive models can improve the accuracy of stock forecasts.

System Architecture / Methodology

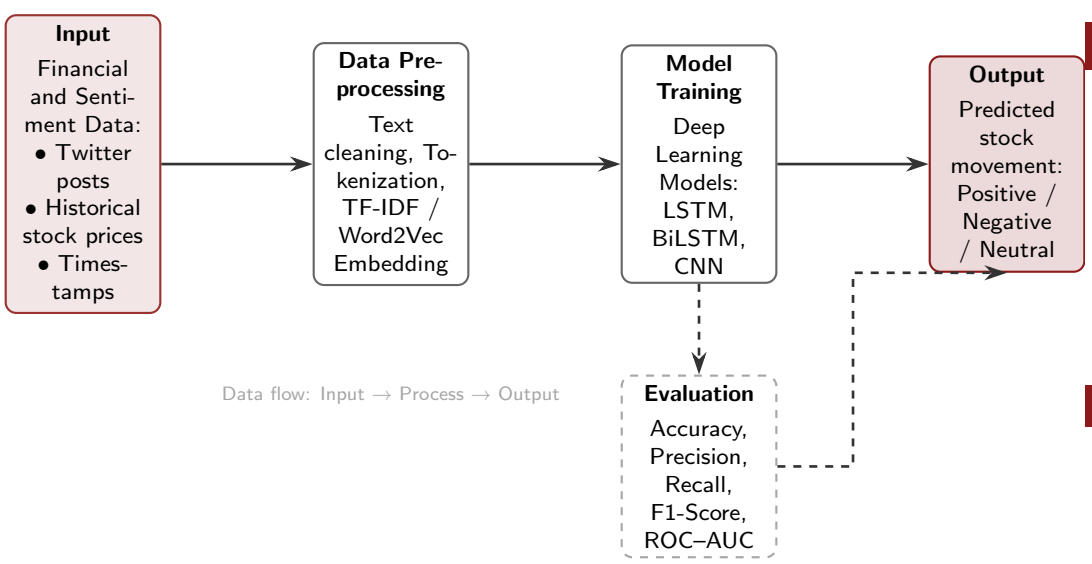


Overall Framework: Input → Preprocessing → Model Training → Evaluation

- **Data Preprocessing:** Tweets were cleaned, tokenized, and converted into numerical embeddings using TF-IDF and Word2Vec for sentiment analysis.
- **Algorithms:** Deep learning models like LSTM, BiLSTM, and CNN were used to predict stock movements from sentiment data.
- **Evaluation Metrics:** Model performance was measured using accuracy, precision, recall, F1-score, and ROC-AUC.

Working of the Model

Input: Historical stock prices and corresponding Twitter posts with timestamps.
Process: Data cleaning → Decision Tree training (C4.5 variants) → Weighting features.
Output: Predicted stock trend or price movement (positive, negative, or neutral) based on social media sentiment.



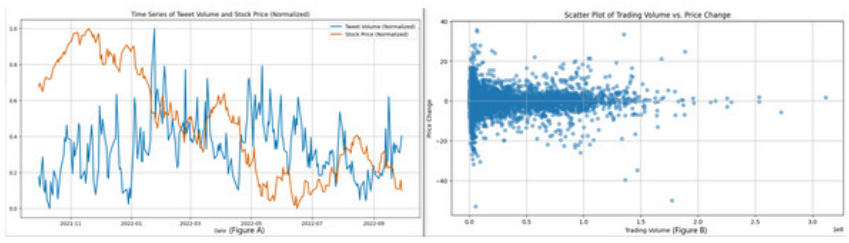
Model Workflow: Input → Process → Output

Dataset Used

- **Original Dataset:** Original research used multiple stocks (Tesla, Apple) .
- **Reproduced Dataset:** *Twitter Stock Market Analysis: Case Study*).
- **Extended Dataset:** *Stock-Market Sentiment Dataset*

Comparative Results and Observations

Model	Accuracy	Precision	Recall	F1
LSTM	82.3%	81.7%	80.9%	81.3%
BiLSTM	84.6%	84.2%	83.9%	84.0%
Hybrid LSTM–CNN	86.8%	86.1%	85.9%	86.0%



(A) Shows the time-series trend of tweet volume and stock price after normalization in the dataset, and (B) is a scatter plot between trading volume and price change.

Observations:

- A positive correlation is visible between tweet volume spikes and short-term stock price fluctuations.
- Higher trading volumes often coincide with larger price changes, indicating market reactions to social sentiment.
- Normalized trends reveal that sentiment activity on Twitter can precede notable movements in stock prices.

Extension on New Dataset

Applied the same methodology .

- **Mean Accuracy:** 52%
- **Precision = 0.52, Recall = 0.52, F1 = 0.52**
- Model performs similarly across both "Up" and "Down" classes, indicating a balanced but low discrimination capability.

Key Insights and Learnings

- Deep learning models, especially Hybrid LSTM–CNN, effectively capture temporal and contextual patterns in financial sentiment data.
- Twitter sentiment shows measurable influence on short-term stock direction.
- Model performance highlights the importance of combining textual and numerical financial features.
- Results validate that hybrid architectures outperform single deep models for stock movement prediction.

Conclusion

The study demonstrates that deep learning models, particularly the Hybrid LSTM–CNN architecture, can effectively predict stock market trends by integrating Twitter sentiment with financial data. The results confirm that social media sentiment significantly influences short-term stock movements, and combining sequential (LSTM) and spatial (CNN) feature learning enhances prediction accuracy. Overall, hybrid deep learning models provide a more reliable framework for sentiment-driven financial forecasting.
References:

Reference

Guo, K., & Xie, H. (2024). *Deep learning in finance: Assessing Twitter sentiment impact and prediction on stocks.* PeerJ Computer Science, 10:e2233. https://doi.org/10.7717/peerj-cs.2233