Hybrid Sentiment-Price Prediction: Leveraging FinBERT and LSTM for Stock Price Movement Forecasting

Author: Reza Tashakkori
reza.tashakkori.001@student.uni.lu
University of Luxembourg

ABSTRACT

Stock price movement prediction is crucial for financial decision-making, yet traditional methods often overlook news and market sentiment, which significantly influence trends. Recent advances in Deep Learning (DL), leveraging Natural Language Processing (NLP) and sequence modeling, provide innovative solutions. This research proposes a hybrid model combining FinBERT for sentiment analysis with a Long Short-Term Memory (LSTM) network to predict price movements, using financial news and stock data from 2020–2024 for Tesla, Apple, Microsoft, and NVIDIA, curated into 33,595 entries. Key contributions include: (1) Fine-tuning FinBERT, achieving 95.24% accuracy and F1 scores (0.94–0.96) across classes, (2) Developing an LSTM model with 85.80% test accuracy and consistent F1 scores (0.85–0.87), and (3) Creating a hybrid model with Test Accuracy 82.47%, F1 scores (>80%), ROC-AUC (~90%), and MCC (0.65), effectively handling financial prediction complexity despite moderate performance. This scalable approach enhances sentiment-driven investment strategies for traders and analysts.

1 INTRODUCTION

Stock price movement prediction is a cornerstone of financial decision-making, guiding traders and analysts through market volatility. Traditional methods often overlook the profound impact of news and market sentiment, which can drive significant price trends. This research addresses this gap by leveraging recent advances in DL, integrating NLP and sequence modeling to combine textual and numerical data effectively. Nowadays, with the widespread use of text data in finance to analyze and predict market trends, financial text mining has become a key pillar of financial technology, enabling innovative solutions like ours. This research proposes a hybrid FinBERT-LSTM model to predict price movements, achieving robust performance with high accuracy and scalability. This work is vital, offering a sentiment-driven tool to enhance forecasting precision in dynamic markets. Its importance extends to countries by boosting economic stability, to companies by informing strategic investments, and to global markets by fostering informed decision-making, ultimately supporting sustainable financial growth.

2 RELATED WORKS

Stock price movement prediction has seen significant advancements through the integration of DL and time-series modeling. Hochreiter and Schmidhuber (1997) introduced LSTM networks, pioneering a method to handle long-term dependencies in sequential data using memory cells and gating mechanisms, laying a foundation for financial time-series analysis [8]. Kraus and Feuerriegel (2017) enhanced this approach by applying LSTM with transfer learning to analyze financial disclosures, focusing on adapting pre-trained models to financial contexts [11]. Nabipour et al. (2020) proposed a hybrid CNN-LSTM model to capture spatial and temporal patterns in stock trends, emphasizing technical indicators over sentiment [13]. Shi et al. (2022) advanced this by developing an attention-based CNN-LSTM-XGBoost hybrid model, integrating multiple architectures to leverage historical stock data for improved prediction [15]. These studies highlight the evolution of sequential and hybrid modeling techniques for financial forecasting.

A parallel strand of research centers on NLP and sentiment analysis within financial contexts. Araci (2019) introduced FinBERT, a pre-trained language model tailored for financial text, adapting transformer architectures to enhance sentiment classification [2]. Liu et al. (2020) built on this by developing FinBERT with advanced tokenization and embeddings to improve financial text mining capabilities [12]. Huang et al. (2022) utilized deep learning to extract key entities and relationships from financial documents, advancing information extraction techniques [7]. Piao et al. (2018) proposed an ensemble approach using CNNs and RNNs for aspect-based sentiment prediction in financial texts, avoiding handcrafted features [14]. Yang et al. (2020) applied FinBERT to analyze corporate reports, focusing on contextual understanding of financial communications [16]. Fjellstrom (2022) explored scalable sentiment classification using an ensemble of LSTM networks for financial time-series data [3]. These efforts underscore the growing role of NLP in interpreting financial narratives.

A third direction involves hybrid and advanced NLP methodologies. Halder (2022) combined FinBERT with LSTM to predict stock prices using news sentiment, emphasizing feature extraction from textual data [6]. Jiang and Zeng (2023) enhanced LSTM forecasting by integrating FinBERT-derived sentiment analysis, proposing methods to refine model performance [9]. Gu et al. (2024) developed a FinBERT-LSTM model for NASDAQ-100 stock prediction, incorporating Benzinga news categories to improve real-time analysis [5]. Aparicio et al. (2024) fine-tuned BioBERT for financial sentiment analysis of biotech press releases, targeting inflection points in stock prices [1]. Kirtac and Germano (2024) leveraged large language models like OPT and FinBERT for sentiment trading, focusing on financial news analysis [10].

The state-of-the-art in this domain aligns closely with the hybrid approaches of Halder (2022), Jiang and Zeng (2023), and Gu et al. (2024), which integrate sentiment and temporal data, as well as the advanced NLP techniques of Aparicio et al. (2024) and Kirtac and Germano (2024), which adapt large language models to financial contexts, mirroring this work's focus on combining textual sentiment with sequential modeling for stock movement prediction.

3 EXPERIMENTS

3.1 Datasets

The FNSPID Dataset (NASDAQ Financial News, provides `Date`, `News Title`, and `Stock Symbol`, with Sentiment Labels (0: Negative, 1: Neutral, 2: Positive) derived via FinBERT [4]). The yFinance API, sourcing stock data from Yahoo Finance, offers `Date`, `Close Price`, and `Price Change` (0: Down, 1: Up). The Final Merged Dataset, with 33,595 rows and 7 columns (`Date`, `Stock Symbol`, `News Title`, `Sentiment Label`, `Close Price`, `Price Change`, `Avg_Sentiment`), was optimized by reducing FinBERT Fine-Tuning runtime from ~12 hours to 1 hour through decreased epoch numbers, batch size, and other settings. Aggregated FNSPID by `Date` and `Stock Symbol`, calculating `Avg_Sentiment` (mean of Sentiment Labels). Merged with yFinance on `Date` and `Stock Symbol` (inner join), ensuring no missing values.

3.2 Model Architectures

Fine-Tuned FinBERT: Built on Google's BERT, FinBERT is a 12-layer Transformer with 768 hidden units, 12 attention heads, and 110M parameters, pre-trained on financial texts for sentiment classification (0: Negative, 1: Neutral, 2: Positive). Fine-tuning optimized weights for our dataset, retaining the same structure.

LSTM Model: A 3-layer LSTM with 512 hidden units processes 3-day sequences of 12 input features (`Avg_Sentiment`, `Close_Price`, `Lag1_Sentiment`, `Lag1_Price`, `Price_Diff`, `Rolling_Sentiment_5d`, `Rolling_Price_Diff_5d`, `News_Volume`, `Rolling_News_Volume_5d`, `Prop_Negative`, `Prop_Neutral`, `Prop_Positive`) for binary classification (Up/Down). Chosen over GRU and CNNs (<75% accuracy) for its ability to capture sequential patterns, prevent vanishing gradients, and suit low-compute datasets compared to Transformers. A 0.3 dropout rate prevents overfitting.

Hybrid Model: Integrates Fine-Tuned FinBERT for sentiment labeling with a pre-trained LSTM. FinBERT generates sentiment features, which are merged with price sequences as input to the LSTM, forming a unified prediction framework without retraining.

3.3 Training Details

Fine-Tuned FinBERT: Final Dataset split into 80% train, 10% validation, and 10% test sets, stratified by company. Used weighted cross-entropy loss, trained for 5 epochs with batch size 32, learning rate 2e-5, 500 warmup steps, 0.01 weight decay, and early stopping (patience=3) to optimize performance and reduce runtime.

LSTM Model: Preprocessed with lagged and rolling features (e.g., 5-day averages), split into 60% train, 20% validation, and 20% test sets, stratified by company. Trained for 50 epochs with batch size 128, learning rate 0.003 using Adam optimizer, learning rate scheduler, weighted cross-entropy loss, 0.01 weight decay, and early stopping (patience=5).

Hybrid Model: No retraining; loaded pre-trained FinBERT and LSTM models. Data preprocessing generated sentiment labels and features, merged with price data, ensuring alignment for evaluation on the test set.

3.4 Evaluation Results

Performance metrics for the models are summarized below, with Fine-Tuned FinBERT abbreviated as FT-FinBERT. The Matthews Correlation Coefficient (MCC), used for binary classification, ranges from +1 (perfect prediction) to 0 (random guessing) to -1 (inverse prediction).

Table 1: Performance Metrics for each Model

Metric	FT-FinBERT	LSTM	Hybrid Model
Epochs	5	50	–
Test Accuracy	95.24%	85.80%	82.47%
Per-Class Accuracy	0: 93.98% 1: 95.01% 2: 96.34%	Down: 83.17% Up: 88.21%	Down: 78.80% Up: 85.76%
F1-Score	0: 94% 1: 95% 2: 96%	Down: 85% Up: 87%	Down: 81% Up: 84%
ROC-AUC	–	–	90%
Runtime	64.47 min	1.19 min	3.72 min
MCC	–	–	0.65

3.5 Replicability

The study is replicable using FNSPID and yFinance datasets with Python, PyTorch, and Hugging Face in a T4-GPU Google Colab environment. Preprocessing includes tokenization (max_length=128) and stratified splitting. Trained models (e.g., `finbert_finetuned.pth`, `lstm_best.pth`, `combined_model.pth`) are saved for reuse. This work supports reproducibility and encourages future extensions, such as incorporating attention mechanisms or expanding to larger datasets and markets.

4 DISCUSSION

FT-FinBERT achieves 95.24% accuracy and F1 scores (0.94–0.96), with validation accuracy rising from 92.38% to 95.15% over 5 epochs, but increasing validation loss after epoch 2 from 0.15653 to 0.29453, suggests mild overfitting. LSTM's 85.80% accuracy and F1 scores (0.85–0.87) reflect strong temporal prediction, with validation accuracy improving from 51.69% to 86.14% and loss dropping to 0.34575, though further training may not improve results. The hybrid FinBERT-LSTM model, with 82.47% accuracy, a robust ROC-AUC of 0.90, and MCC of 0.65, advances sentiment-driven stock prediction, though it shows a slight accuracy drop from LSTM's 85.80%, indicating sentiment integration challenges. Future enhancements could explore advanced data integration techniques, better handling of class imbalances, and real-time adaptability, roadening its applicability in volatile financial environments.

REFERENCES

[1] Valentina Aparicio, Daniel Gordon, Sebastian G. Huayamares, and Yuhuai Luo. 2024. BioFinBERT: Finetuning Large Language Models (LLMs) to Analyze Sentiment of Press Releases and Financial Text Around Inflection Points of Biotech Stocks. *arXiv* (2024). https://doi.org/10.48550/arXiv.2401.11011

[2] Dogu Araci. 2019. FinBERT: Financial Sentiment Analysis with Pre-trained Language Models. *arXiv* (2019). https://doi.org/10.48550/arXiv.1908.10063

[3] Carmina Fjellstrom. 2022. Long Short-Term Memory Neural Network for Financial Time Series. *arXiv* (2022). https://doi.org/10.48550/arXiv.2201.08218

[4] FNSPID Dataset. 2023. FNSPID-filtered-nasdaq-100: NASDAQ Financial News and Sentiment Data. https://huggingface.co/datasets/benstaf/FNSPID-filtered-nasdaq-100

[5] Wen jun Gu, Yi hao Zhong, and Shi zun Li. 2024. Predicting Stock Prices with FinBERT-LSTM: Integrating News Sentiment Analysis. *ACM Digital Library* (2024). https://doi.org/10.1145/3694860.3694870

[6] Shayan Halder. 2022. FinBERT-LSTM: Deep Learning Based Stock Price Prediction Using News Sentiment Analysis. *arXiv* (2022). https://doi.org/10.48550/arXiv.2211.07392

[7] Allen H.Huang, Hui Wang, and Yi Yang. 2022. FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research* (2022). https://doi.org/10.1111/1911-3846.12832

[8] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* (1997). https://doi.org/10.1162/neco.1997.9.8.1735

[9] Tingsong Jiang and Qingyun Zeng. 2023. Financial sentiment analysis using FinBERT with application in predicting stock movement. *arXiv* (2023). https://doi.org/10.48550/arXiv.2306.02136

[10] Kemal Kirtac and Guido Germano. 2024. Sentiment trading with large language models. *arXiv* (2024). https://doi.org/10.1016/j.frl.2024.105227

[11] Mathias Kraus and Stefan Feuerriegel. 2017. Decision Support from Financial Disclosures with Deep Neural Networks and Transfer Learning. *Decision Support Systems* (2017). https://doi.org/10.1016/j.dss.2017.10.001

[12] Zhuang Liu, Degen Huang, and Kaiyu Huang. 2020. FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining. *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (2020). https://doi.org/10.24963/ijcai.2020/622

[13] Mojtaba Nabipour, Pardis Nayyeri, and Hamed Jabani. 2020. Deep Learning for Stock Market Prediction. *Entropy* (2020). https://doi.org/10.3390/e22080840

[14] Guangyuan Piao and John G. Breslin. 2018. Financial Aspect and Sentiment Predictions with Deep Neural Networks: An Ensemble Approach. *ACM Digital Library* (2018). https://doi.org/10.1145/3184558.3191829

[15] Zhuangwei Shi, Yang Hu, Guangliang Mo, and Jian Wu. 2022. Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction. *arXiv* (2022). https://doi.org/10.48550/arXiv.2204.02623

[16] Yi Yang, Mark Christopher Siy UY, and Allen Huang. 2020. FinBERT: A Pretrained Language Model for Financial Communications. *arXiv* (2020). https://doi.org/10.48550/arXiv.2006.08097

A EDA OF FINAL MERGED DATASET

The following figures visualize key characteristics of the final merged dataset, illustrating sentiment distributions, news article counts, price changes, and class weight balancing across companies from 2020 to 2024.

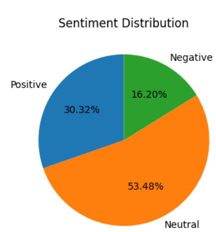

Figure 1: Overall Sentiment Distribution

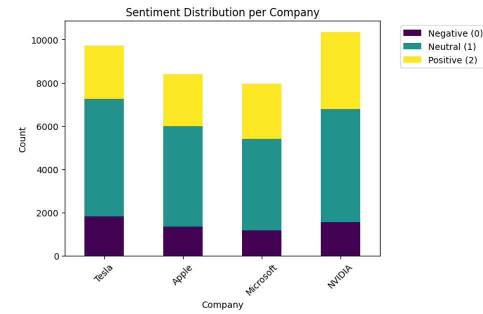

Figure 2: Sentiment Distribution per Company

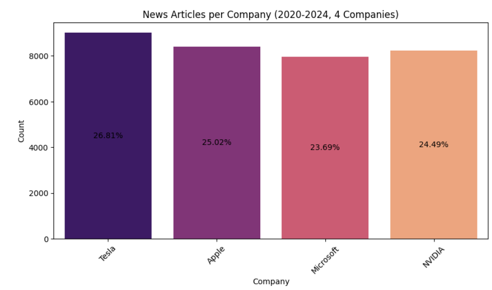

Figure 3: News Articles per Company (2020–2024)

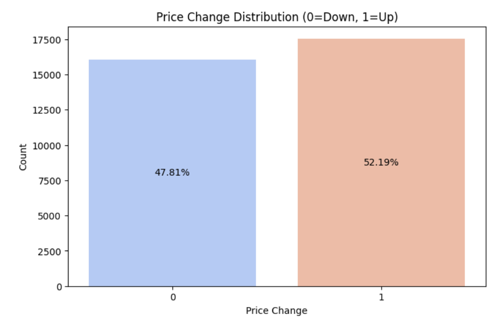

Figure 4: Price Change Distribution (0=Down, 1=Up)

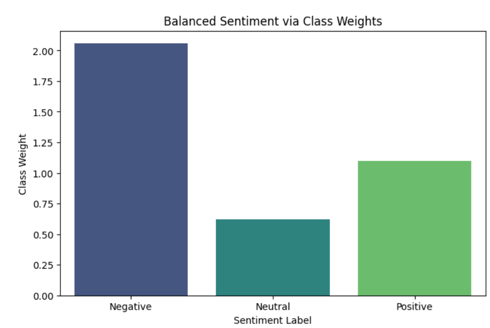

Figure 5: Balanced Sentiment via Class Weights