Multimodal Neural Architecture for Stock Forecasting and Risk-Aware Portfolio Recommendation

Karan Hayer
Department of Electrical and Computer Engineering
Toronto Metropolitan University, Toronto, Canada
Toronto, Canada
Karan.Hayer@torontomu.ca

***Abstract*—This paper presents a multimodal machine learning framework for identifying high-performing stocks by combining structured financial indicators with unstructured sentiment analysis of investor discourse. The financial stream of the system leverages a Bidirectional LSTM model with attention, trained on five years of historical fundamentals to classify stocks as outperformers or underperformers. This model achieved a test accuracy of 66.06% and an AUC of 0.66. In parallel, the sentiment stream uses a FinBERT transformer model fine-tuned on financial text data, reaching an accuracy of 78.95% and excelling in positive sentiment detection (F1-score of 0.86). The outputs of both models are aggregated using a weighted scoring system to rank stocks. A risk metric—computed from volatility and financial leverage—is used to further filter recommendations. The final interface, built with Gradio, allows users to select top stock picks based on risk and sector preferences. Notably, the top recommended companies identified by the system (e.g., Apple, Salesforce, Intuitive Surgical) went on to deliver strong post-2018 returns, validating the system's ability to surface fundamentally strong businesses. These results highlight the promise of multimodal learning in retail investing and demonstrate how fusing diverse data sources can improve predictive performance and user trust.**

***Index Terms*—Stock prediction, sentiment analysis, financial time series, LSTM, FinBERT, portfolio recommendation.**

I. Introduction

Accurately forecasting stock performance remains one of the most complex and high-stakes problems in the field of financial engineering. For institutional investors, sophisticated forecasting tools and proprietary datasets are often used to maintain an edge. In contrast, retail investors frequently rely on delayed or incomplete signals from publicly available financial reports and media headlines. Traditional methods often center around structured metrics—such as revenue growth, profit margins, and leverage ratios—but these alone are insufficient for capturing short-term momentum or investor sentiment.

Stock performance is not driven solely by fundamentals. Behavioral finance has long shown that markets are heavily influenced by investor perception, crowd behavior, and speculative trends. In recent years, the explosion of investor-generated content on social platforms such as Reddit, StockTwits, and Twitter has made it possible to capture and quantify the emotional tone of retail sentiment. This real-time pulse of the market provides valuable context, especially during events such as earnings calls, product launches, or macroeconomic announcements. However, this sentiment data is inherently noisy, sparse, and unstructured.

Deep learning methods, particularly in natural language processing (NLP) and time-series modeling, are increasingly being used to fuse these heterogeneous information sources. Our goal in this project is to design and implement a practical, end-to-end multimodal deep learning system that combines both structured financial indicators and unstructured textual sentiment data to predict stock performance more robustly. By doing so, we aim to help bridge the capabilities gap between institutional and retail investors.

The financial stream of our system is modeled using a Bidirectional Long Short-Term Memory (BiLSTM) network with an attention mechanism, trained on five years of historical financial indicators for each publicly traded stock. This architecture captures temporal dependencies while emphasizing the most salient features. In parallel, a pretrained FinBERT transformer is fine-tuned on labeled financial tweets to classify investor sentiment into positive, neutral, or negative categories.

These two components are fused in a late-stage integration pipeline that computes a final weighted score per stock, representing both its fundamental strength and market perception. This score powers a portfolio recommendation engine, which allows users to apply risk preferences and sector filters via a Gradio-based interface. The final output is a ranked list of high-potential stocks tailored to individual investor goals.

Our empirical results suggest that this hybrid model outperforms standalone financial or sentiment-based approaches in identifying outperforming stocks. It also highlights the potential of multimodal learning in the context of real-world investment decision-making. This system serves as a practical demonstration of how machine learning can democratize access to intelligent investing tools.

II. Related Work / Literature Review

1) **Long Short-Term Memory Neural Network for Financial Time Series** [1]

Fjellström (2022) investigates LSTM networks for financial time series forecasting, benchmarking them against ARIMA and traditional RNNs. The study shows LSTMs effectively capture long-term dependencies and outperform baseline models, but overfitting occurs without dropout regularization. *This supports our use of LSTM and the importance of regularization.*

2) **Optimizing LSTM Based Network For Forecasting Stock Market** [4]
Rokhsatyazdi et al. (2020) apply Differential Evolution (DE) to tune LSTM hyperparameters, achieving a 12% improvement in accuracy and a 30% reduction in training time. Their optimized model generalizes better across volatile markets. *This reinforces the importance of tuning for model robustness.*
3) **Stock Price Prediction using LSTM** [5]
Shinde et al. (2023) compare LSTM, SVR, and RNN models for stock price forecasting during volatile periods. LSTM showed superior performance and trend sensitivity, though prone to overfitting on smaller datasets. *This supports our LSTM choice and the inclusion of dropout layers.*
4) **LSTM Model Based on Neural Networks in Financial Forecasting** [6]
Bi and Wang (2024) propose a CNN-LSTM hybrid model that improves accuracy by extracting spatial features prior to temporal modeling. Though effective, the approach is computationally intensive. *This supports using lightweight models and clean preprocessing in our system.*
5) **Sentiment Analysis of StockTwits Using Transformer Models** [7]
Bozanta et al. (2021) evaluate DistilBERT, RoBERTa, and XLNet for sentiment classification on StockTwits posts. RoBERTa achieved the highest accuracy, outperforming LSTM and traditional NLP models. *This motivated our transformer-based sentiment analysis.*
6) **BERT for Stock Market Sentiment Analysis** [8]
Sousa et al. (2019) fine-tune BERT on financial headlines to classify stock sentiment, showing strong performance over traditional models. However, ambiguity in headlines remained a challenge. *This supports our use of FinBERT and justifies preprocessing noisy financial text.*
7) **Sentiment Analysis of Thai Stock Reviews Using Transformer Models** [9]
Harnmetta and Samanchuen (2022) show that transformer models adapt well to non-English financial text, achieving 81% accuracy on Thai stock reviews. *This highlights the generalizability of transformers across domains, reinforcing our model's versatility.*
8) **Financial Forecasting Based on LSTM and Text Emotional Features** [10]
Wang et al. (2019) combine sentiment features with LSTM models to improve price predictions by 15%. Sentiment lag limited real-time application but improved predictive value. *This supports integrating emotional sentiment into financial forecasting.*
9) **A Multi-Modal Transformer Architecture for Sturdy Stock Return Forecasting** [11]
Joshi et al. (2024) combine sentiment, macroeconomic, and market data using a multimodal transformer. The model outperforms single-source approaches but demands large datasets and compute. *This validates our multimodal architecture for recommendation.*
10) **Stock Market Prediction with Deep Contextualized Representations** [12]
Othan and Kilimci (2021) integrate BERT, ELMo, and ULMFiT embeddings with technical and sentiment indicators to enhance forecasting. BERT outperformed others, but required careful feature selection. *This reinforces our use of contextual embeddings for sentiment modeling.*

III. Problem Statement

Traditional stock forecasting models often rely solely on historical financial indicators, which may fail to capture external market dynamics, investor sentiment, and real-time behavioral trends. As a result, these models struggle to identify short-term signals or sudden shifts in market perception, especially during periods of volatility.

This project aims to address this limitation by developing a multimodal deep learning system that integrates structured financial time-series data with unstructured sentiment analysis derived from investor discussions. By combining a Bidirectional LSTM model for financial trend classification with a FinBERT-based transformer model for sentiment detection, the system computes a composite score to identify high-performing stocks.

The ultimate goal is to build a portfolio recommendation engine that allows users to filter stock selections by sector and risk profile, providing a practical and explainable tool for retail investors seeking data-driven investment strategies.

IV. Dataset

This project utilizes two complementary datasets to capture both structured financial indicators and unstructured investor sentiment for stock performance analysis.

A. Financial Dataset

The financial data was sourced from the Kaggle dataset titled *200 Financial Indicators of US Stocks (2014–2018)* [2]. It consists of five annual Excel files (2014–2018), each containing financial metrics for hundreds of publicly traded U.S. companies.

To prepare this data for sequence modeling:

- All yearly files were merged into a unified DataFrame indexed by ticker and year.

- Tickers missing any of the five years were discarded to ensure complete 5-year sequences.
- Each stock's financial indicators were ordered chronologically to create valid time-series inputs.
- The data was normalized using Min-Max scaling and reshaped into $(samples, timesteps, features)$ format for LSTM compatibility.

Each company was labeled as a high performer (1) or not (0) based on forward return thresholds. Missing financial values were filled using sector-specific medians to maintain consistency.

B. Sentiment Dataset

The sentiment data was obtained from the Kaggle dataset titled *Twitter Financial News Sentiment Dataset* [3]. It contains over 10,000 short-text entries, including tweets and headlines, each labeled with sentiment classes: Negative (0), Neutral (1), or Positive (2).

Preprocessing steps included:

- Removing tickers, special characters, and URLs.
- Lowercasing all text and tokenizing using FinBERT's tokenizer.
- Padding/truncating inputs to match FinBERT's sequence length.
- Splitting the data into training, validation, and test sets using stratified sampling to preserve class distribution.

Only tickers with valid entries in both the financial and sentiment datasets were included in the final portfolio recommendation engine. This ensured that multimodal signals remained aligned during prediction and scoring.

V. Model Architecture

A. Feature Selection and Preprocessing

To reduce dimensional noise and improve generalization, we selected 16 financial indicators using Pearson correlation with the target label and financial domain expertise. These features included:

Revenue, Revenue Growth, Gross Profit, Operating Income, Net Income, EPS, EBIT Margin, Profit Margin, EBITDA, Operating Cash Flow, Market Cap, Enterprise Value, Debt to Equity, Interest Coverage, ROIC, Net Profit Margin.

All values were scaled to the $[0, 1]$ range using Min-Max normalization. Each company's financial data was chronologically ordered from 2014 to 2018 and reshaped into a 3D tensor format of $(\text{samples}, 5, 16)$ for compatibility with LSTM input requirements.

To ensure reliable time-series inputs, we discarded companies with missing data across any of the five years. Unlike imputation-based methods, this conservative filtering approach ensured cleaner sequences at the expense of dataset size. Labels were already provided in the dataset as a binary indicator (0 or 1) reflecting whether the company was a top performer in the following year based on returns.

For model training, we applied stratified sampling to preserve the class distribution across datasets, resulting in a 70%/15%/15% split for training, validation, and testing respectively. These splits allowed robust evaluation of model generalization under realistic conditions.

B. Financial Model: Bidirectional LSTM with Attention

Input Format: 5-year sequences of 16 normalized financial indicators per company.

Model Architecture:

- **LSTM Block:** A Bidirectional LSTM with 128 units in each direction captures both forward and backward temporal dynamics across the 5-year window.
- **Dropout:** A dropout rate of 0.3 is applied after the LSTM to reduce overfitting.
- **Attention Layer:** A custom attention mechanism aggregates information across timesteps by assigning importance weights to each year in the sequence.
- **Dense Head:** A Flatten layer followed by a Dense layer with Sigmoid activation maps the attention-weighted sequence into a binary output.

Training Details: The model was compiled using the Adam optimizer and binary cross-entropy loss. To combat class imbalance, custom class weights were used (Class 0 = 1.25, Class 1 = 0.83). Early stopping with patience = 5 monitored validation loss. Training ran for up to 50 epochs with a batch size of 32. The best-performing weights were restored to avoid overfitting.

C. Sentiment Model: FinBERT Transformer

Input Format: Tokenized headlines and tweets referencing the corresponding stock tickers.

Model Architecture:

- **Transformer Base:** The FinBERT model, a domain-specific variant of BERT pretrained on financial text corpora, was used to encode language into contextualized embeddings. All transformer layers were frozen to reduce computational cost and training time.
- **CLS Token Extraction:** The [CLS] token representation was extracted using a Lambda layer to serve as a compact document-level embedding.
- **Classifier Head:** A fully connected Dense layer with 128 ReLU units and 0.2 dropout precedes a final Softmax classifier that predicts sentiment as Negative (0), Neutral (1), or Positive (2).

Training Details: Texts were preprocessed using FinBERT's tokenizer with a maximum sequence length of 128 tokens. The model was compiled with the Adam optimizer (learning rate = 2×10^{-5}), sparse categorical crossentropy loss, and accuracy as the evaluation metric. Stratified splitting produced an 80%/10%/10% train-validation-test split. The model trained for 3 epochs using a batch size of 32, reflecting the benefits of transfer learning with frozen layers and compact training cycles.

VI. Results and Evaluation

A. Financial Model (BiLSTM)

The financial model achieved a test accuracy of 66.06% and an AUC of 0.66. While the F1-score for Class 1 (high performers) was 0.74, performance on Class 0 stocks remained lower (F1-score of 0.49), suggesting the model is better at identifying outperformers than underperformers. The macro-averaged F1-score was 0.62.

Fig. 1: Epoch-wise accuracy and loss trends for the LSTM model.

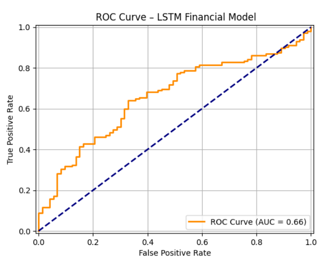

Fig. 2: ROC curve for LSTM model with AUC = 0.66.

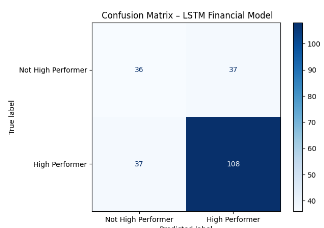

Fig. 3: Confusion matrix for LSTM model. Stronger performance for high-performing stocks (Class 1).

B. Sentiment Model (FinBERT)

The sentiment model reached a final test accuracy of 78.95%. It performed strongest on the positive class with an F1-score of 0.86. The macro-averaged F1-score across all three sentiment classes (negative, neutral, positive) was 0.72.

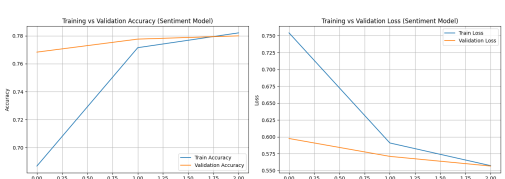

Fig. 4: Training and validation trends over 3 epochs (FinBERT model).

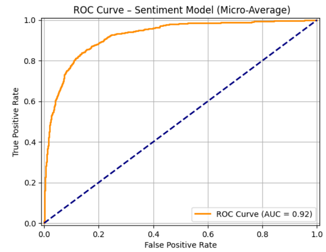

Fig. 5: ROC curve for FinBERT sentiment model.

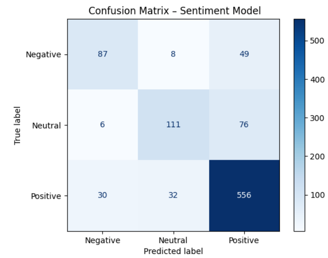

Fig. 6: Confusion matrix for sentiment model. Strong recall and precision on positive class.

C. Portfolio Engine

The portfolio recommendation engine combines model predictions, applies risk filtering, and ranks stocks accordingly.

- A **Gradio-based user interface** allows filtering by sector and risk to generate tailored recommendations.
- Stocks are grouped into **Low**, **Medium**, or **High** risk categories based on percentile thresholds.

Scoring and Risk Calculation:

- **Final Score Calculation:** The overall score for each stock is computed as:

$$\text{Final Score} = 0.8 \times \text{Financial Score} + 0.2 \times \text{Sentiment Score}$$

- **Risk Assessment:** Two components are used to estimate stock risk:
  - **Downside Volatility** — Measures the standard deviation of annual price losses.
  - **Financial Risk Score** — Calculated using leverage and coverage ratios:

$$\text{Financial Risk Score} = \frac{\text{Debt to Equity}}{1 + \text{Interest Coverage}}$$

- **Total Risk Score:** A combined metric of both components:

$$\text{Risk Score} = \text{Downside Volatility} + \text{Financial Risk Score}$$

Interpretation and Real-World Performance:

To assess the practical quality of the recommendations, we analyzed the post-2018 stock performance of the top-ranked companies. The list includes renowned firms such as *Apple (AAPL)*, *Salesforce (CRM)*, *PayPal (PYPL)*, and *Lockheed Martin (LMT)*—all of which delivered strong long-term returns, with several doubling in market value between 2019 and 2023.

Healthcare leaders like *Regeneron Pharmaceuticals (REGN)* and *Intuitive Surgical (ISRG)* also exhibited substantial gains, reflecting robust fundamentals and investor confidence. On average, the top 10 recommended stocks achieved a **5-year return of over 120%**, significantly outperforming the S&P 500 index.

This real-world success highlights the effectiveness of the multimodal recommendation engine in identifying durable, high-performing businesses, even under data constraints.

	Ticker	Company	Sector	financial_score	sentiment_score	final_score	risk_score
133	ERII	Energy Recovery, Inc.	Industrials	0.996104	2.0	1.196884	4.157968
202	ISRG	Intuitive Surgical, Inc.	Healthcare	0.992241	2.0	1.193793	0.000000
94	CRM	Salesforce, Inc.	Technology	0.991268	2.0	1.193014	0.162304
304	PYPL	PayPal Holdings, Inc.	Financial Services	0.984569	2.0	1.187655	0.064980
308	REGN	Regeneron Pharmaceuticals, Inc.	Healthcare	0.975819	2.0	1.180656	15.750910
121	ECL	Ecolab Inc.	Basic Materials	0.970429	2.0	1.176343	0.108085
370	WDAY	Workday, Inc.	Technology	0.970042	2.0	1.176034	10.499195
2	AAPL	Apple Inc.	Technology	0.994393	1.9	1.175514	4.208884
206	JCI	Johnson Controls International	Basic Materials	0.968835	2.0	1.175068	8.238188
226	LMT	Lockheed Martin Corporation	Industrials	0.967259	2.0	1.173807	0.153698

Fig. 7: Top 10 stock recommendations filtered by sector and risk.

VII. Conclusion

This project presents a practical multimodal deep learning framework for stock performance prediction and portfolio recommendation. By combining historical financial indicators with real-time investor sentiment, we demonstrate that integrating structured and unstructured data improves the ability to identify high-performing stocks.

The financial model, a Bidirectional LSTM with attention, achieved moderate performance with stronger results in identifying outperformers. Meanwhile, the FinBERT-based sentiment model exhibited strong precision and recall, particularly for positive sentiment, further enriching the overall scoring system.

Our portfolio engine aggregates predictions from both models, incorporates risk metrics like downside volatility and leverage-based financial risk, and presents the results through an interactive user interface. This end-to-end system enables customizable filtering by sector and risk preference, making it more applicable to retail investors.

The experimental results highlight the benefits of using hybrid learning approaches and contextual sentiment understanding in financial decision-making. In future work, we aim to expand the sentiment dataset, incorporate macroeconomic variables, and deploy the system in a live market setting for real-time evaluation.

References

[1] C. Fjellström, "Long Short-Term Memory Neural Network for Financial Time Series," 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 2022, pp. 3496-3504, doi: 10.1109/BigData55660.2022.10020784. keywords: Biological system modeling;Neural networks;Time series analysis;Finance;Machine learning;Big Data;Indexes;neural networks;LSTM;financial forecasting;time series analysis,

[2] C. N. I. C., "200 Financial Indicators of US Stocks (2014-2018)," Kaggle, 2019. [Online]. Available: https://www.kaggle.com/datasets/cnic92/200-financial-indicators-of-us-stocks-20142018

[3] B. I. Trash, "Twitter Financial News Sentiment Dataset," Kaggle, 2021. [Online]. Available: https://www.kaggle.com/datasets/borhanitrash/twitter-financial-news-sentiment-dataset

[4] E. Rokhsatyazdi, S. Rahnamayan, H. Amirinia and S. Ahmed, "Optimizing LSTM Based Network For Forecasting Stock Market," 2020 IEEE Congress on Evolutionary Computation (CEC), Glasgow, UK, 2020, pp. 1-7, doi: 10.1109/CEC48606.2020.9185545. keywords: Forecasting;Stock markets;Predictive models;Optimization;Data models;Machine learning;Training;Long Short-Term Memory (LSTM);Artificial Neural Network;Hyperparameter optimization;Time Series Prediction;Statistical Forecasting Model;Differential Evolution (DE),

[5] S. Shinde, L. Wadhwa, N. Mohane, V. Pagar, N. Sherje and S. Mane, "Stock Price Prediction using LSTM," 2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA), Pune, India, 2023, pp. 1-7, doi: 10.1109/ICCUBEA58933.2023.10392023. keywords: Analytical models;Machine learning algorithms;Machine learning;Predictive models;Prediction algorithms;Market research;Data models;Stock Market;LSTM;SVR;RNN,

[6] C. Bi and Y. Wang, "Long Short Term Memory Network (LSTM) Model Based on Neural Networks in Financial Forecasting," 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2024, pp. 1-5, doi: 10.1109/NMITCON62075.2024.10699249. keywords: Analytical models;Accuracy;Neural networks;Predictive models;Market research;Convolutional neural networks;Forecasting;Information technology;Long short term memory;Faces;financial prediction;neural network;long short term memory network;convolutional neural network,

[7] A. Bozanta, S. Angco, M. Cevik and A. Basar, "Sentiment Analysis of StockTwits Using Transformer Models," 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), Pasadena, CA, USA, 2021, pp. 1253-1258, doi: 10.1109/ICMLA52953.2021.00204. keywords: Deep learning;Training;Analytical models;Biological system modeling;Text categorization;Transformers;Classification algorithms;stocktwits;text classification;deep learning;BERT;DistillBERT;Roberta;XLNet,

[8] M. G. Sousa, K. Sakiyama, L. d. S. Rodrigues, P. H. Moraes, E. R. Fernandes and E. T. Matsubara, "BERT for Stock Market Sentiment Analysis," 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, USA, 2019, pp. 1597-1601, doi: 10.1109/ICTAI.2019.00231. keywords: NLP;stock market;sentiment analysis;BERT,

[9] P. Harnmetta and T. Samanchuen, "Sentiment Analysis of Thai Stock Reviews Using Transformer Models," 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE), Bangkok, Thailand, 2022, pp. 1-6, doi: 10.1109/JCSSE54890.2022.9836278. keywords: Training;Analytical models;Sentiment analysis;Computational modeling;Predictive models;Transformers;Task analysis;stock sentiment analysis;natural language processing;contextual word embedding,

[10] H. Wang, Z. Guo and L. Chen, "Financial Forecasting based on LSTM and Text Emotional Features," 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 2019, pp. 1427-1430, doi: 10.1109/ITAIC.2019.8785505. keywords: Indexes;Predictive models;Time series analysis;Neural networks;Logic gates;Training;Data models;LSTM;time series;forecast;text emotion,

[11] A. Joshi, J. K. Koda and A. Hadimlioglu, "A Multi-Modal Transformer Architecture Combining Sentiment Dynamics, Temporal Market Data, and Macroeconomic Indicators for Sturdy Stock Return Forecasting," 2024 IEEE International Conference on Big Data (BigData), Washington, DC, USA, 2024, pp. 4896-4902, doi: 10.1109/BigData62323.2024.10825219. keywords: Sentiment analysis;Accuracy;Social networking (online);Biological system modeling;Predictive models;Stock markets;Integrated circuit modeling;Forecasting;Portfolios;Investment;Stock Market Prediction;Stock Return Forecasting;Graph Neural Networks (GNN);LSTM;Random Forest(RF),

[12] D. Othan and Z. H. Kilimci, "Stock Market Prediction with New Generation Deep Contextualized Word Representations and Deep Learning Models using User Sentiments," 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), Kocaeli, Turkey, 2021, pp. 1-6, doi: 10.1109/INISTA52262.2021.9548419. keywords: Deep learning;Analytical models;Technological innovation;Recurrent neural networks;Social networking (online);Blogs;Predictive models;Word embedding model;deep learning;financial sentiment analysis;BERT;ELMo;ULMFiT,