HAW HAMBURG

Forecasting Financial Futures: LSTM Approach to Tech, Gold & Bitcoin

Prateek Kalra / Lars Henske / Janice Kohrs / Da Yeon Kang

1. Introduction & Goal

We predict short-term price changes for the **top 10 US tech stocks, Gold,** and **Bitcoin** using LSTM, a deep learning model suited for time-series data. By combining historical price data with news headline sentiment **(Aug 18, 2023 - Mar 18, 2024)**, we achieved more accurate predictions than using price data alone.

2. What is LSTM?

LSTM (Long Short-Term Memory) is a **neural network** designed for time-series data. It remembers important past information and filters out noise using special gates, making it ideal for predicting future trends like stock prices.

3. Which stocks did we analyze?

- **Top 10 US stocks -Trading Volume**

Apple (AAPL)	Microsoft (MSFT)
Amazon (AMZN)	Alphabet (GOOGL)
Nvidia (NVDA)	Tesla (TSLA)
Palantir (PLTR)	Meta (META)
Uber Technologies (UBER)	Robinhood Markets (HOOD)

- **Bitcoin & Gold**

4. Methodology

5 classes based on 5-day future return:
- 0: Strong drop (≤ −2%)
- 1: Small drop (−2% to −0.5%)
- 2: Neutral (−0.5% to +0.5%)
- 3: Small rise (+0.5% to +2%)
- 4: Strong rise (> +2%)

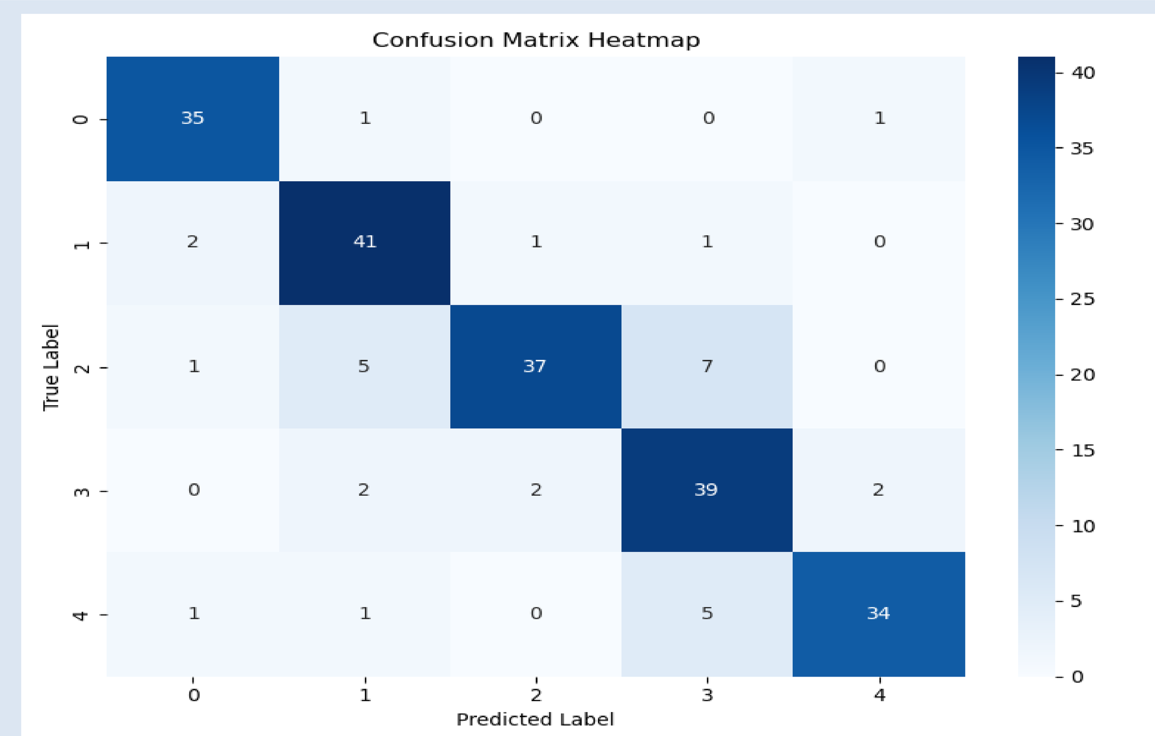
Input: 40-day window of standardized technical indicators (e.g., MACD, RSI, MA, OBV)
Target: Return from day 40 to 45
Model: Stacked LSTM + dropout, batch norm, softmax

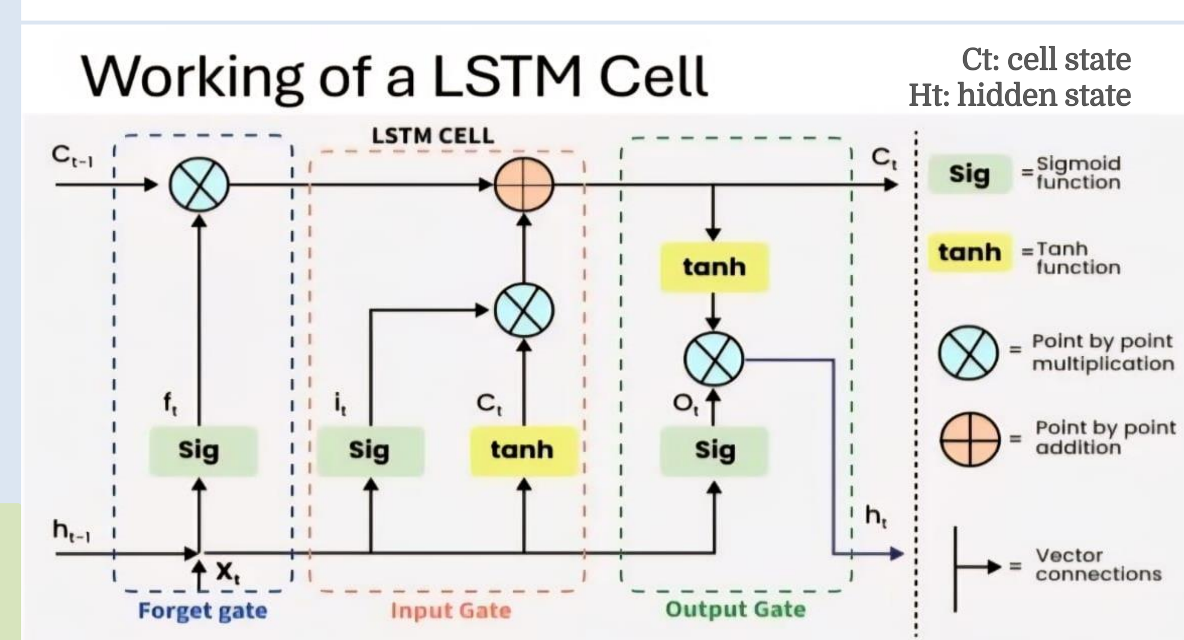
Training:
- Early stopping (patience=15, restore best weights)
- Class weighting to boost extremes (classes 0 & 4)

5. Results

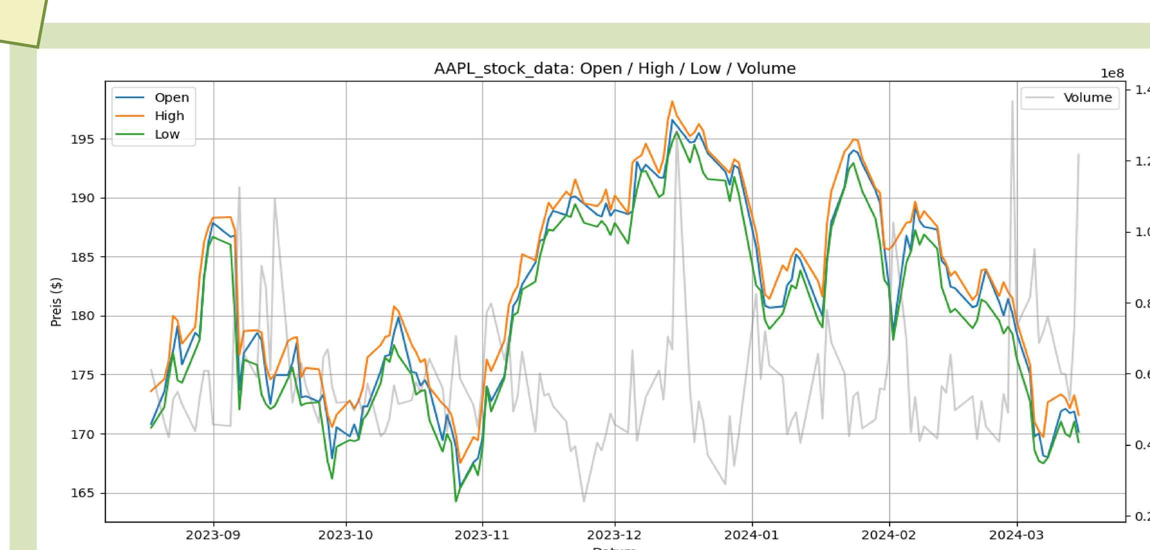
(1) Model Performance Summary:

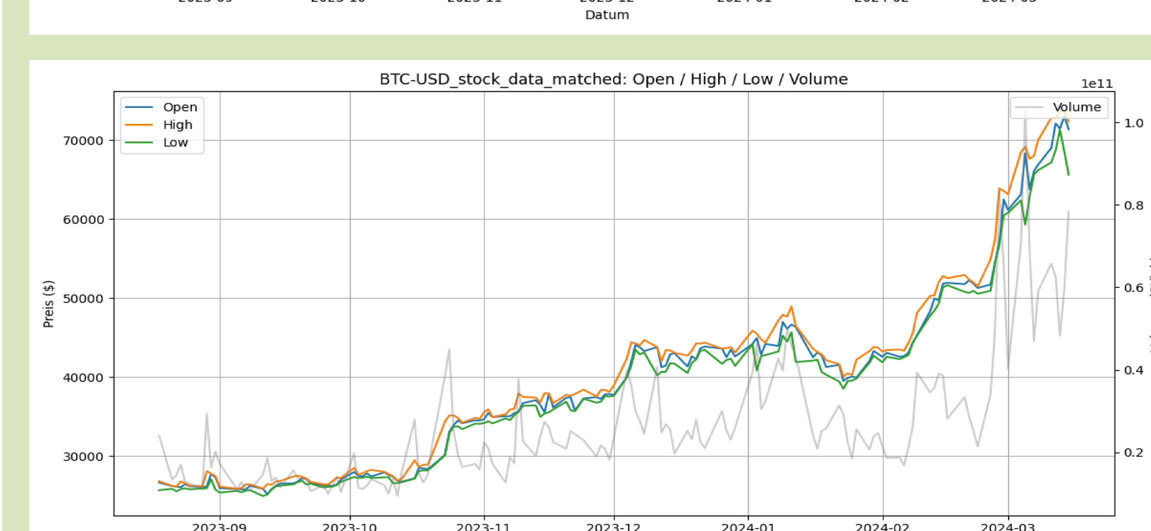
The LSTM model shows strong overall accuracy, with the confusion matrix indicating most predictions align with true labels. Misclassifications are minimal and mostly between adjacent classes. Class 0 and 1 exceed 90% accuracy, while Class 2 is lowest (~75%) due to volatility. Overall, the model captures temporal stock patterns well with balanced, robust performance.





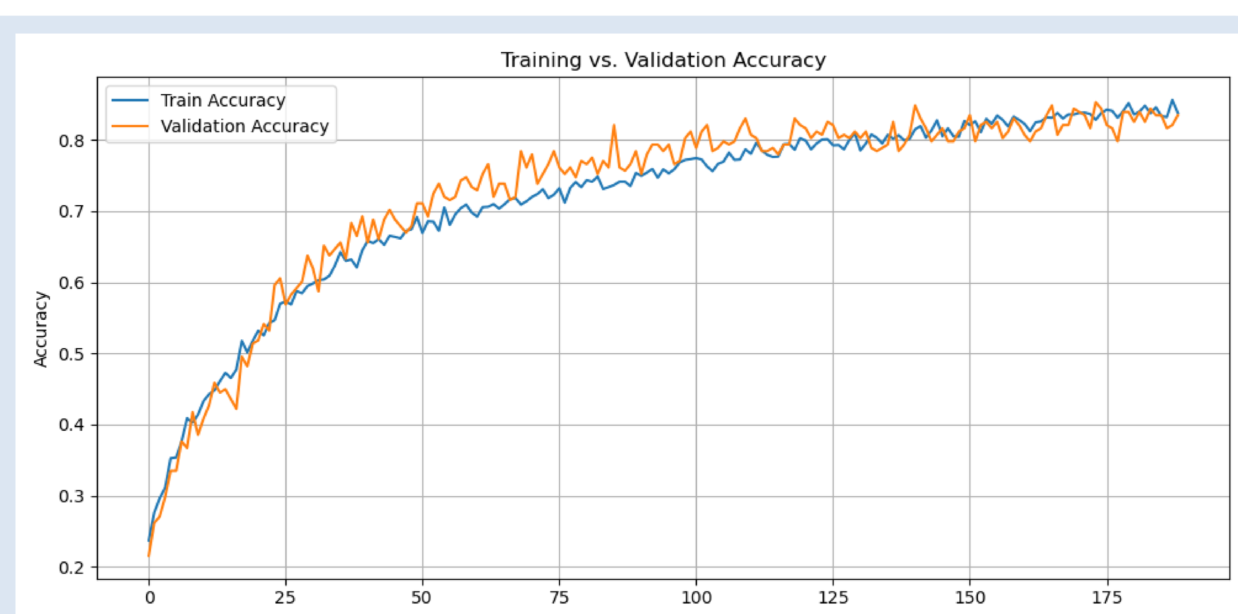
*Stock data Open/High/Low/Volume ex. Apple, BitCoin

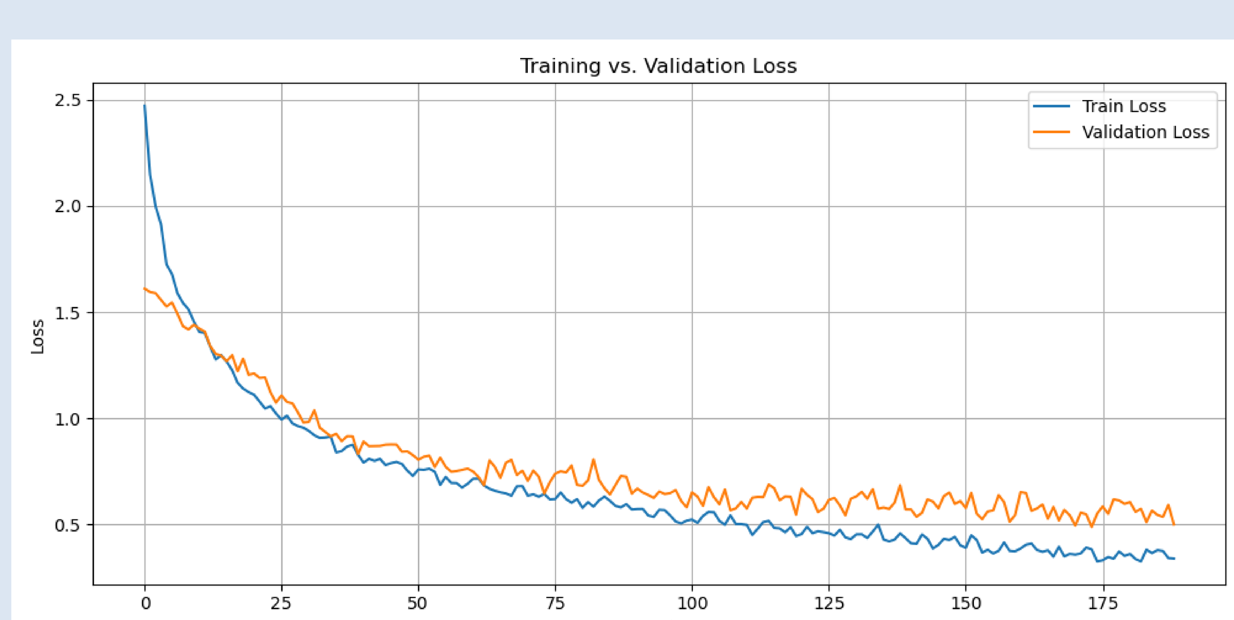




(2) Training vs. Validation Loss:

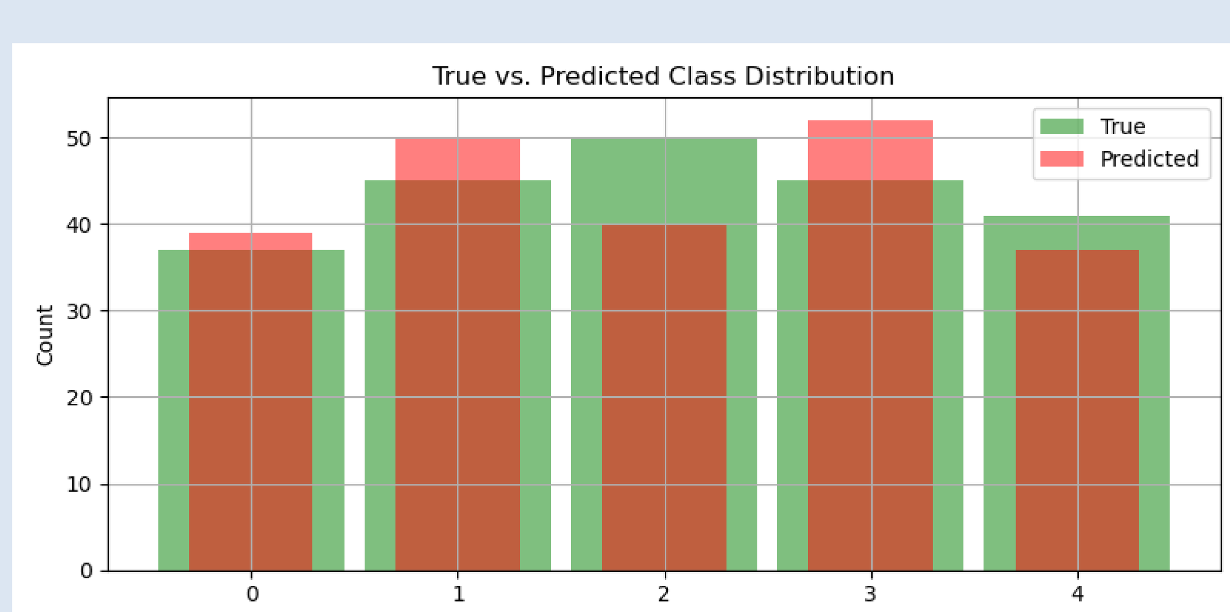
Both training and validation losses drop sharply during the first 80–100 epochs, indicating effective learning. Training loss continues to decline, while validation loss plateaus around epoch 100, signaling mild overfitting. Dropout and early stopping (patience=15) help prevent further overfitting, maintaining model generalization.





(3) True vs. Predicted Class Distribution:

The predicted class counts closely match the true distribution, indicating balanced predictions. Slight overprediction of Class 1 suggests it's easier to model due to smoother patterns, while Classes 2 and 4 are slightly underpredicted, likely due to volatility. Overall, the model avoids class bias—a key strength for real-world financial applications.



6. Conclusion

Our LSTM model performs well despite market volatility, effectively capturing short-term trends and managing bias-variance. It's especially accurate for high-cap, high-volume stocks and works best when combined with domain expertise. Future enhancements include sentiment analysis, attention mechanisms, and multimodal data integration.