

PROBLEM

Stock price prediction is challenging due to market volatility and non-linear patterns. Traditional models struggle to capture temporal dependencies in financial time series.

OBJECTIVES

- Predict next-day stock closing price
- Evaluate model accuracy using historical data
- Compare LSTM(Long Short-Term Memory) vs baseline models

LIMITATIONS

- Cannot predict sudden market shocks
- Short-term direction accuracy is low
- Performance decreases for longer prediction horizons

KEY FINDINGS

- LSTM captures overall stock price trends
- Achieves ~5% average prediction error
- Direction prediction remains difficult (~52%)
- Accuracy decreases for longer prediction horizons

FUTURE WORK

- Multi-day prediction (5–10 days)
- Sentiment analysis (news data)
- Real-time deployment

SYSTEM ARCHITECTURE

Pipeline for stock price prediction using LSTM:



METHODOLOGY

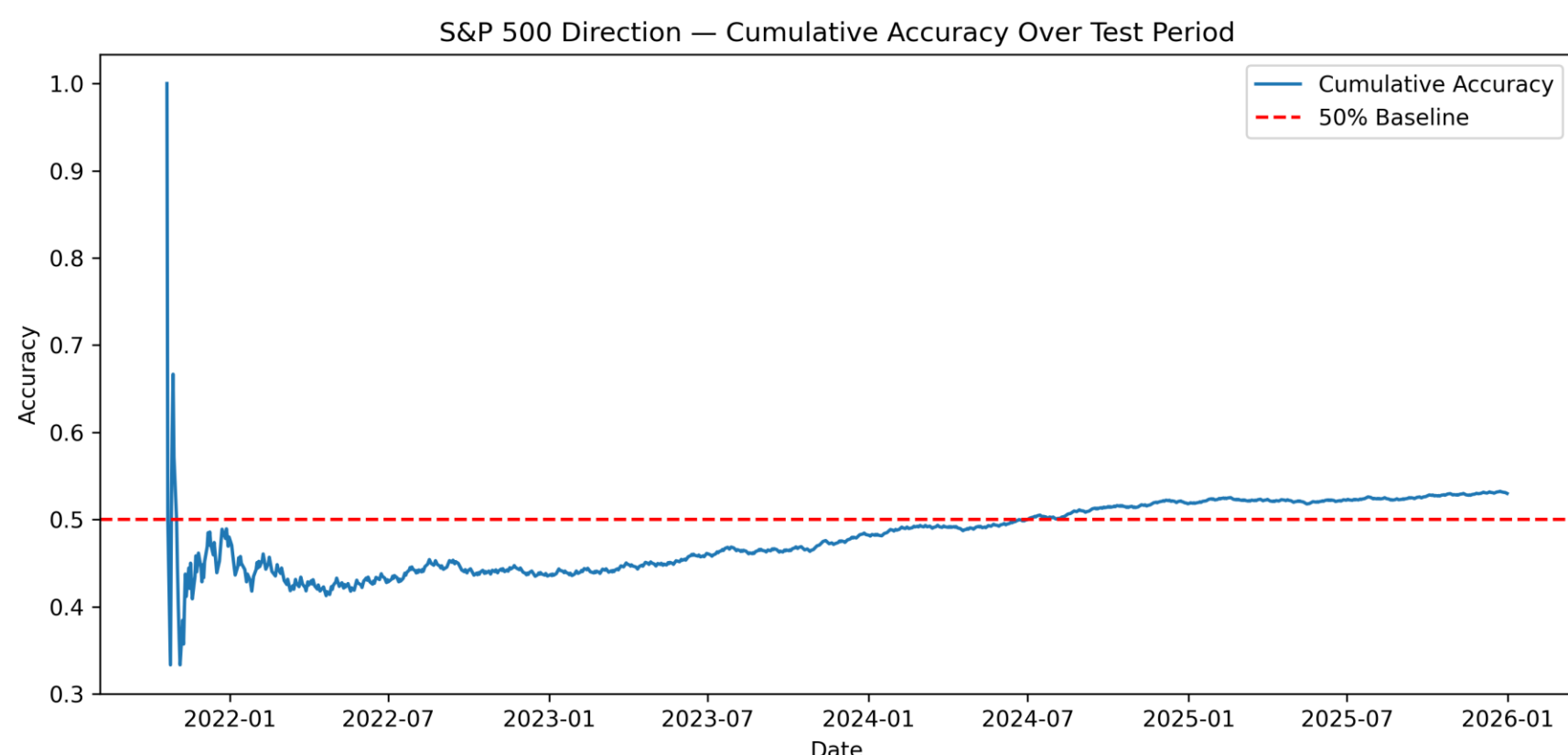
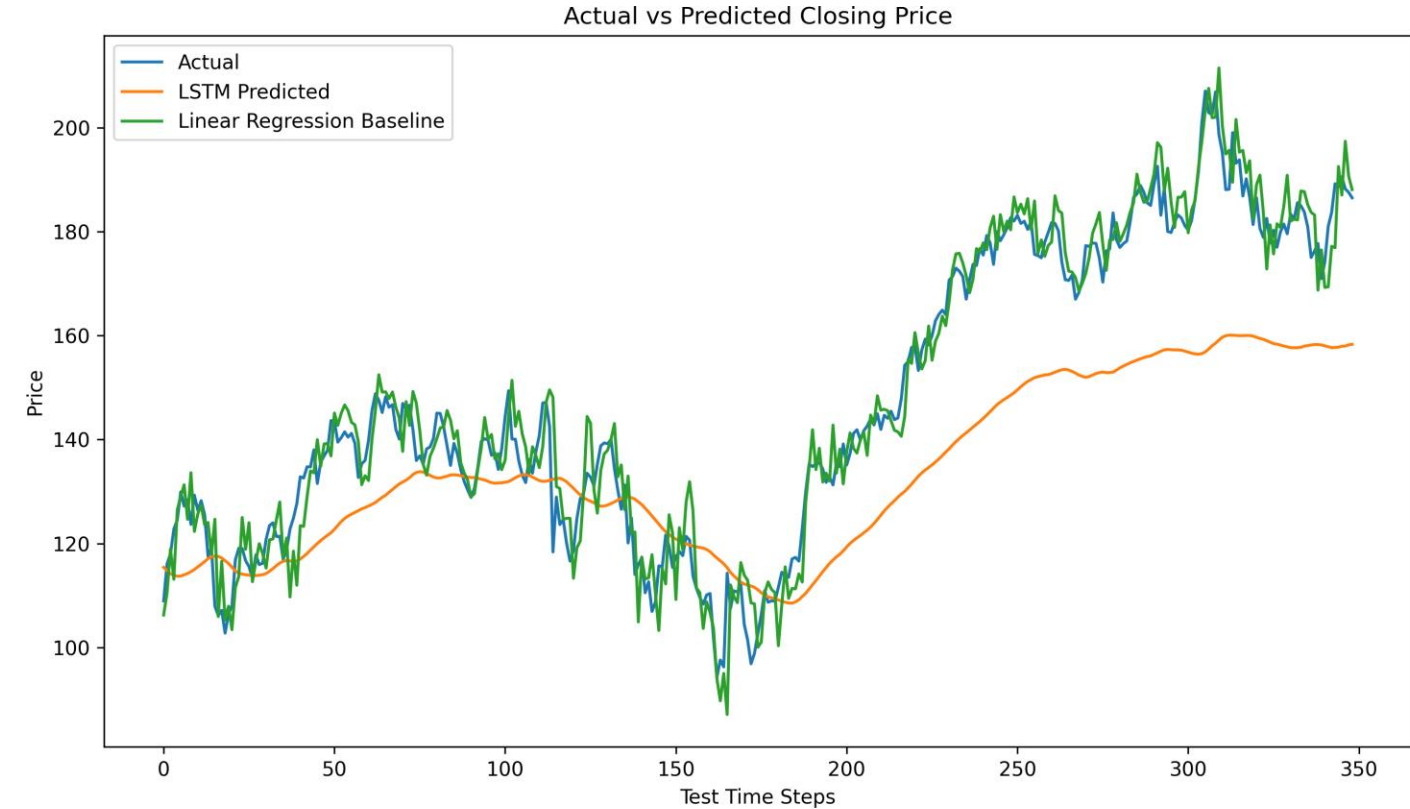
Data & Features

- AAPL, NVDA stock data (2019–2026)
- Moving averages (MA_5, MA_10, MA_20)
- RSI, MACD, volatility
- Market indices (S&P 500, NASDAQ)

Model & Training

- LSTM Neural Network (128 → 64 layers)
- Sequence length: 30 days
- Train/test split: 80/20
- Loss: Mean Squared Error
- Early stopping applied

RESULTS



Conclusion:

- Linear regression performs strongly for next-day prediction due to high short-term autocorrelation in stock prices
- LSTM captures broader temporal patterns and produces smoother, more generalized predictions
- Directional prediction remains challenging, with accuracy only slightly above random (~52%)

BACKTESTING & MODEL PERFORMANCE

Backtesting evaluates model performance by simulating predictions on unseen historical data over time.

Key Insights:

- LSTM model performance improved by increasing model capacity (from 32→64 to 128→64 units), reducing prediction error and improving trend capture.
- Directional accuracy remains close to random (~52%), highlighting the difficulty of predicting short-term market movement.
- Linear regression performs strongly due to short-term autocorrelation, while LSTM provides smoother trend modeling.

Model	RMSE	MAE	R ²
Linear Regression	~5	~3.5	~0.95
LSTM(Initial)	~25	~20	~0.3
LSTM(Improved)	~18	~14	~0.45
LSTM(Final)	15	12	0.58

RMSE and MAE measure prediction error, MAPE shows percentage error, and R² indicates how well the model explains price variation.