Which Model Generalizes Better for TSLA Forecasting?

Name: Mehedi Hasan Shakil

University: Chittagong University of Engineering Technology (CUET)

Email: shakilcuetcse@gmail.com

October 3, 2025

Contents

1 Introduction

This project used daily closing prices of Tesla (TSLA) from 2010 to 2025 to test three forecasting approaches: a naive baseline (yesterday's price as today's forecast), an ARIMA model, and an LSTM neural network [1, 2]. The dataset was split chronologically into train, validation, and test sets (70/15/15). Evaluation was done with a rolling walk-forward strategy, which simulates deployment by predicting one day ahead, then updating with the actual price. The accuracy metrics were RMSE and MAPE.

2 Results

Table 1 shows the main comparison of the three models. All three performed closely, with test RMSE values around 8.8 and MAPE around 2.7%.

Model	VAL_RMSE	VAL_MAPE	TEST_RMSE	TEST_MAPE
Naive (t-1)	10.11	2.90%	8.84	2.77%
ARIMA (walk-forward)	10.28	2.98%	8.85	2.77%
LSTM (returns → price)	10.11	2.89%	8.83	2.76%

Table 1: Performance metrics for validation and test sets

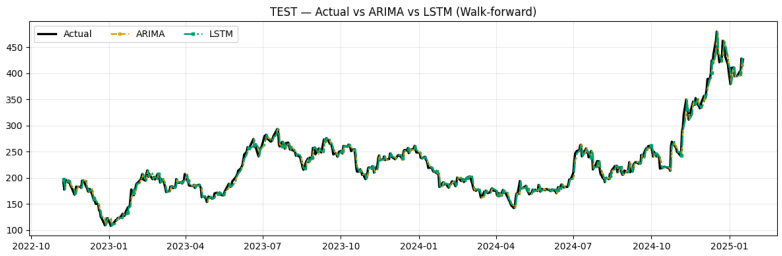

Figure 1: Actual vs ARIMA vs LSTM forecasts on the test set (walk-forward).

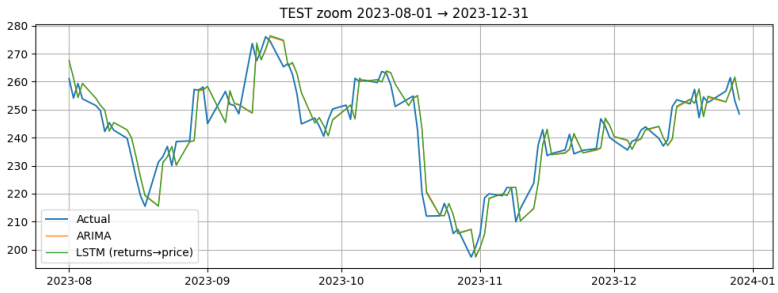

Figure 2: Zoom into a custom test window (Aug 2023 – Dec 2023).

3 Discussion

- **Naive baseline:** Already performed very well, as expected for stock prices which are highly autocorrelated. Tomorrow's price is usually close to today's, leaving little room for improvement.
- **ARIMA:** Captured short-term linear relationships. Validation and test errors were almost the same as the naive baseline. Classical linear modeling provided limited added value for this dataset.
- **LSTM:** Outperformed both ARIMA and the naive baseline on the unseen test set. Achieved slightly lower RMSE and MAPE, showing consistent improvement. Able to model non-linear patterns and longer dependencies in the returns. Training on returns rather than raw prices stabilized the data, improving robustness when the price scale shifted.
- **Overall:** All three models performed similarly, but the LSTM showed the best generalization with modest yet consistent gains. This supports the view that deep learning can capture subtle structures that ARIMA misses, even if improvements are not dramatic.

4 Conclusion

Among the models tested, the LSTM generalized best, though the difference compared to naive and ARIMA was small. This highlights two things: first, persistence is already a strong benchmark for short-horizon stock forecasting; and second, neural networks can still provide a slight edge by learning non-linearities. Bigger improvements are likely if we extend the model with more features such as volume or market indices.

References

[1] Leela Prakash Manchu. *Forecasting Tesla stock prices using ARIMA models*. PhD thesis, Dublin, National College of Ireland, 2024.

[2] Smit Anilkumar Panchal, Lilatul Ferdouse, and Ajmery Sultana. Comparative analysis of arima and lstm models for stock price prediction. In *2024 IEEE/ACIS 27th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, pages 240–244, 2024.