Stock Market Prediction Using LSTM, GRU, and Transformer Models on Tesla Stock Data (2015–2024)

1st Asmae Mouradi
Department of Computer Science
Wichita State University
Wichita, USA

2nd Renuka Mohanta
Department of Computer Science
Wichita State University
Wichita, USA

3rd Dr. Lokesh Das
Department of Computer Science
Wichita State University
Wichita, USA

***Abstract*—In this project we compare deep learning models (LSTM, GRU, combined LSTM+GRU, and Transformer) for forecasting Tesla (TSLA) daily closing prices using data from 2015–2024. We train univariate and multivariate variants and evaluate 1, 7, and 15 days ahead forecasts. The best performance is obtained with the GRU second approach (medium configuration) for one-step-ahead forecasting (H=1), achieving RMSE = 7.3000, MSE = 53.2900, MAE = 5.4166, and $R^2 = 0.9609$ on the test set. The code is available in the following GitHub repository.**

***Index Terms*—Stock price forecasting, LSTM, GRU, Transformer, Time series analysis, Tesla stock**

I. Introduction

In this project, we investigate several deep learning architectures for forecasting the daily closing price of Tesla (TSLA) stock from 2015 to 2024. Using historical OHLCAV data, we implement and compare univariate and multivariate versions of Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), a combined LSTM+GRU model, and an encoder-only Transformer. All models are evaluated in three capacity settings (small, medium, large) for one-step and multi-step horizons $H \in \{1, 7, 15\}$ using RMSE, MSE, MAE and R^2 as evaluation metrics. Small, medium, and large refer to configurations with progressively larger lookback windows, more layers, and more hidden units. On the univariate closing-price series, the best GRU configuration achieves next-day prediction accuracy of approximately RMSE $\approx$ 7 USD with $R^2 \approx 0.96$, outperforming the corresponding LSTM, LSTM+GRU and Transformer variants. Extending the input to a multivariate OHLCAV representation generally improves the stability of multi-step forecasts; in particular, a small multivariate LSTM attains $R^2 > 0.94$ for horizons up to 15 days.

Overall, the results indicate that moderately sized GRU and LSTM architectures provide the best trade-off between accuracy and complexity for forecasting highly volatile single-stock time series such as TSLA.

II. Literature Review

Deep learning methods, particularly recurrent architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have become a standard approach for modeling financial time series. Their gating mechanisms help to mitigate vanishing gradients and allow the models to capture long-range temporal dependencies that arise in noisy and highly non-linear stock price data. In the context of stock market prediction, the literature includes both *univariate* formulations, where only one price series typically the closing price is modeled, and *multivariate* formulations, which incorporate additional variables such as open, high, low, and technical indicators.

A. Univariate LSTM models for stock prices

Prabhu and Kundargi [1] investigate univariate LSTM models for forecasting Apple stock prices using ten years of historical data. Their input consists solely of the daily closing price, scaled with a Min–Max normalisation, and they compare a vanilla single-layer LSTM with a stacked LSTM. Using a sliding window of 60 past days to predict the next day's close, they evaluate performance via the root mean squared error (RMSE) and show that the stacked LSTM substantially outperforms the vanilla configuration (best RMSE ≈ 0.012 versus 0.302), highlighting that deeper architectures can improve predictive accuracy even when only a single feature is used.

Srivastava and Mishra [2] also focus on a univariate deep learning formulation but for Tesla (TSLA) stock. They compare several baselines including moving average, linear regression and ARIMA against an RNN-LSTM model trained on historical prices from Yahoo Finance.They found that LSTM substantially reduces prediction error compared with classical statistical and shallow machine learning models, reinforcing the advantage of sequence models for highly volatile stocks.

Chi [3] extends this line of work by analysing Tesla closing prices over a longer horizon (2013–2023) and explicitly comparing pure LSTM, pure GRU and a hybrid LSTM–GRU architecture. All three models are trained with comparable hyperparameters (200 epochs, 32 hidden units, MSE loss and the Adam optimiser) on normalised daily price data. The study evaluates performance using RMSE, MSE, MAE and R^2 for both training and test sets, and reports that GRU tends to perform better for short-term test forecasts, while the hybrid model shows slightly stronger ability to capture longer-term trends.

B. Multivariate LSTM/GRU models with technical indicators

While the above works demonstrate that accurate forecasts are possible even with a single price series, other authors have argued that incorporating additional market information can further improve performance. Sivadasan et al. [4] propose a multivariate deep learning framework for stock market forecasting based on LSTM and GRU architectures. In a first phase, they use day-wise open, high, low and close (OHLC) values for several stocks (e.g. Intel, Indian Oil, NTPC, Citigroup) and show that both LSTM and GRU outperform existing benchmark models in terms of standard error metrics such as MAE, RMSE, MAPE and R^2.

Overall, the literature indicates that: (i) univariate LSTM and GRU models using only the closing price can already achieve competitive performance on individual stocks such as AAPL and TSLA [1]–[3]; and (ii) multivariate formulations that integrate OHLC data and technical indicators tend to further reduce forecast errors across diverse equities [4].

III. Dataset

The Tesla (TSLA) dataset is publically available was downloaded from Kaggle and originally sourced from Yahoo Finance. It contains 2,843 records from 2015–2024, each representing one trading day. The dataset includes the following fields see table I:

TABLE I
Stock Market Data Description

Column Name	Description
Date	Trading day
Open	Opening stock price
High	Highest price of the day
Low	Lowest price of the day
Close	Closing stock price
Adjusted Close	Adjusted close after splits/dividends
Volume	Shares traded during the day

We start by exploring the dataset; no missing values were found.

IV. Motivation

The motivation of this paper is to compare modern deep learning models for accurately forecasting a highly volatile stock (Tesla) and determine whether more complex architectures outperform simpler RNN-based approaches in practice.

V. Architecture

A. LSTM

The architecture of the LSTM is presented in figure 1, the LSTM model takes as input a window of L past days with F features per day ($F = 1$ for the univariate Close series, $F = 5$ for the multivariate OHLCAV case). This sequence is passed through one or more LSTM layers, which read the days one by one in time and, via their internal gates and memory state, learn how information from earlier days influences future values. A dropout layer is applied to the LSTM output to regularize the model and reduce overfitting. The final hidden representation is fed to a dense layer with a single unit (Dense(1)) to produce a one-step-ahead prediction of the TSLA closing price. The same one-step model is then applied recursively to obtain forecasts for multiple horizons $H \in \{1, 7, 15\}$.

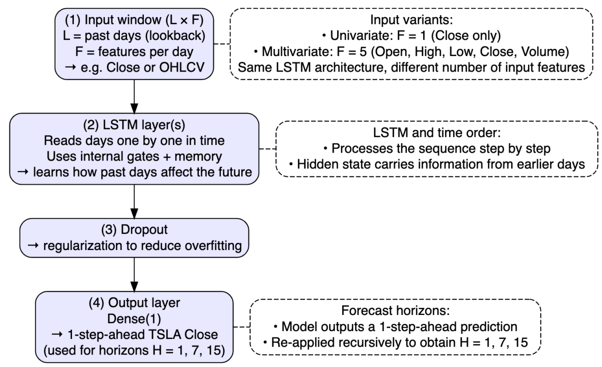

Fig. 1. LSTM architecture

B. GRU

The architecture for the GRU is presented in figure 2, the GRU model uses the same input and output structure as the LSTM. Unlike LSTMs, GRUs have a single hidden state and use only two gates (update and reset) to control how past information is combined with the current input. The same one-step GRU model is applied recursively to obtain forecasts for horizons $H \in \{1, 7, 15\}$.

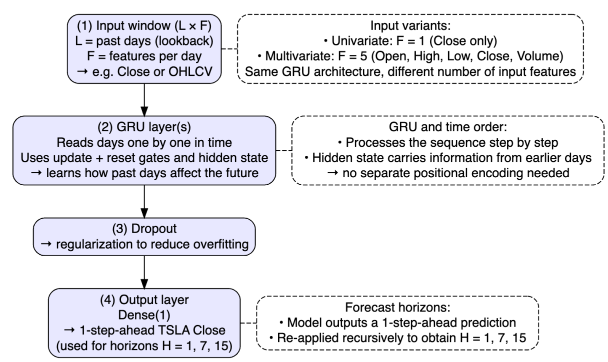

Fig. 2. GRU architecture

C. Transformer

We adopt an encoder-only Transformer architecture as shown in 3 because our forecasting task maps a fixed-length input window to a single scalar output (next-day closing price). The model takes a window of L past days with F features per day as input ($F = 1$ for the univariate Close series, $F = 5$ for the multivariate OHLCAV case). A dense layer projects each time step to a fixed-size embedding of dimension d_{model}, and positional encodings are added to encode the temporal order of the days. The resulting sequence is processed by N Transformer encoder layers (multi-head self-attention and feed-forward sublayers). A global average pooling over the time dimension produces a single summary vector, which is passed through a final dense layer with one unit to predict the next-day TSLA closing price.

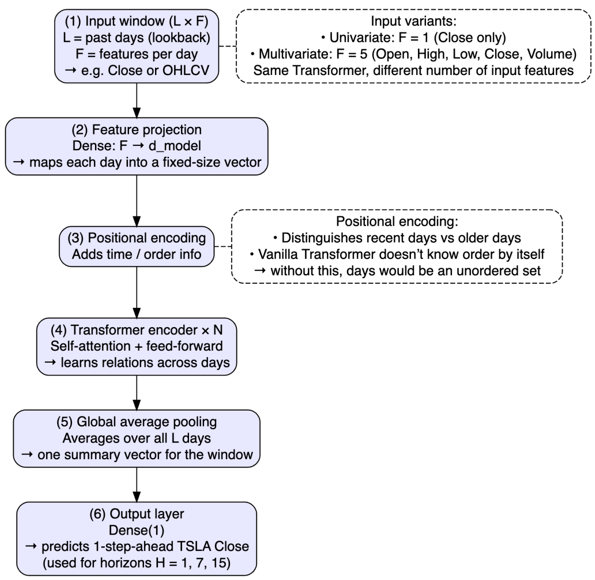

Fig. 3. Transformer Architecture

VI. Training Procedure

A. Univariate time-series forecasting

In this project we treat the TSLA daily closing price as a univariate time series and use deep learning models (LSTM, GRU, Transformer, LSTM+GRU) for next-step and multi-step forecasting. The dataset is split into training (80%), validation (10%), and test (10%) subsets. For each model (LSTM, GRU, Transformer and combined LSTM+GRU), we define three versions: *small*, *medium* and *large*, which correspond to increasing lookback window, number of layers and number of units (model capacity). We will present three different approaches for each model with various sizes to compare their performance. Each model has lookback window, number of layers and units, dropout, batch size, learning rate, and loss are shown in Table II, Table III, Table IV, Table V.

1) LSTM: The hyperparameters of the three univariate LSTM variants are summarized in Table II.

TABLE II
LSTM Univariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
LSTM-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
LSTM-2 (medium)	60	2	64	0.1	5	1×10^{-3}	MSE
LSTM-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

2) Transformer: The hyperparameters of the three univariate Transformer variants are summarized in Table III.

TABLE III
Transformer Univariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss	Attention Head	Feed Forward Dimension
Transformer-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE	2	64
Transformer-2 (medium)	60	2	64	0.1	5	1×10^{-3}	MSE	4	128
Transformer-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE	8	256

3) GRU: The hyperparameters of the three Univariate variants of GRU are summarized in Table VIII

TABLE IV
GRU Univariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
GRU-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
GRU-2 (medium)	60	2	64	0.1	16	5×10^{-3}	MSE
GRU-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

4) Combined LSTM and GRU: The hyperparameters of the three Univariate variants of Combined are summarized in Table IX

TABLE V
LSTM+GRU Univariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
LSTM+GRU-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
LSTM+GRU-2 (medium)	60	2	64	0.1	16	5×10^{-3}	MSE
LSTM+GRU-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

B. Multivariate time-series forecasting

1) LSTM: The hyperparameters of the three multivariate LSTM variants are summarized in Table VI.

TABLE VI
LSTM Multivariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
LSTM-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
LSTM-2 (medium)	60	2	64	0.1	5	1×10^{-3}	MSE
LSTM-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

2) Transformer: The hyperparameters of the three multivariate transformer variants are summarized in Table VII

TABLE VII
Transformer Multivariate experimental

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss	Attention Head	Feed Forward Dimension
Transformer-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE	2	64
Transformer-2 (medium)	60	2	64	0.1	5	1×10^{-3}	MSE	4	128
Transformer-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE	8	256

3) GRU: The hyperparameters of the three multivariate variants of GRU are summarized in Table VIII

TABLE VIII
GRU MULTIVARIATE EXPERIMENTAL

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
GRU-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
GRU-2 (medium)	60	2	64	0.1	16	5×10^{-3}	MSE
GRU-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

4) Combined GRU and LSTM: The hyperparameters of the three multivariate variants of Combined are summarized in Table IX

TABLE IX
LSTM+GRU MULTIVARIATE EXPERIMENTAL

Experiment	Lookback window	Layers	Units	Dropout	Batch	LR	Loss
LSTM+GRU-1 (small)	30	1	32	0.1	16	1×10^{-3}	MSE
LSTM+GRU-2 (medium)	60	2	64	0.1	16	5×10^{-3}	MSE
LSTM+GRU-3 (large)	90	3	128	0.3	16	5×10^{-4}	MSE

VII. RESULTS

A. Evaluation Metrics

We compute standard regression metrics on the test set (and analogously on the training set):

- Root Mean Squared Error (RMSE):It measures the average squared difference between predicted and actual prices. It penalizes large errors strongly, useful for detecting big prediction mistakes.
- Mean Squared Error (MSE):It is the square root of MSE and error is expressed in actual stock price units (USD).Shows the typical size of prediction error.
- Mean Absolute Error (MAE):It is the average of absolute errors between predicted and real values.Gives a clear picture of day-to-day prediction accuracy.
- Coefficient of determination R^2:Measures how well the model explains the variability of Tesla stock prices.Higher R² means better trend prediction.

We repeat this pipeline for the four models and within each approach meaning small, baseline, large defined in the hyperparameter tables, and compare their performance across horizons to study the impact of lookback length, depth, capacity, and regularization on forecasting accuracy.

Model performance was evaluated using:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \quad RMSE = \sqrt{MSE}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$$

B. Key Results

Once trained, the model is evaluated both on one-step-ahead predictions and on multi-step horizons $H \in \{1, 7, 15\}$.

1) Univariate Time-Series Forecasting: In this section, we present the key results for univariate time-series forecasting using various models: LSTM, GRU, Transformer, and Combined LSTM + GRU.

- **LSTM**: X, XI represent results for the univariate LSTM model for three different approaches.

TABLE X
LSTM-1 (SMALL) TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	8.6463	74.7591	6.4038	0.9451
LSTM(H=7)	18.0483	325.7401	14.7767	0.7520
LSTM(H=15)	24.8783	618.9279	21.0602	0.4910

TABLE XI
LSTM-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	10.4860	109.9565	9.1134	0.9901
LSTM(H=7)	30.4755	928.7580	27.7597	0.9165
LSTM(H=15)	59.5377	3544.7375	55.6339	0.6820

TABLE XII
LSTM-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	7.9337	62.9434	6.7378	0.9944
LSTM(H=7)	32.8131	1076.6998	30.2773	0.9041
LSTM(H=15)	74.6331	5570.1055	69.9084	0.5050

TABLE XIII
LSTM-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	7.3898	54.6093	5.5797	0.9599
LSTM(H=7)	20.7395	430.1277	17.0593	0.6725
LSTM(H=15)	33.4689	1120.1646	27.4833	0.0788

TABLE XIV
LSTM-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	20.5869	423.8213	17.1496	0.6889
LSTM(H=7)	26.8608	721.5002	22.5002	0.4506
LSTM(H=15)	34.6140	1198.1288	28.9087	0.0147

TABLE XV
LSTM-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	42.4789	1804.4576	39.6083	0.8405
LSTM(H=7)	64.8981	4211.7617	60.4754	0.6284
LSTM(H=15)	106.0973	11256.6406	98.1391	0.0095

- **Transformer**: Results for the Transformer model.

TABLE XVI
TRANSFORMER-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	42.5830	1813.3137	40.6898	0.8381
Transformer(H=7)	66.2124	4384.0767	64.0056	0.6094
Transformer(H=15)	87.8257	7713.3477	84.8342	0.3145

TABLE XVII
TRANSFORMER-1 (SMALL) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	30.0875	905.2595	27.3798	0.3356
Transformer(H=7)	66.5508	4429.0103	61.4326	-2.3723
Transformer(H=15)	104.4163	10902.7607	98.6539	-7.9661

TABLE XVIII
TRANSFORMER-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	169.5036	28731.4785	156.4608	-1.5649
Transformer(H=7)	169.3833	28690.6895	156.3074	-1.5564
Transformer(H=15)	169.1087	28597.7559	155.9975	-1.5416

TABLE XIX
TRANSFORMER-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	37.1860	1382.7961	33.4354	-0.0149
Transformer(H=7)	36.4136	1325.9528	32.7527	-0.0096
Transformer(H=15)	34.9799	1223.5906	31.6477	-0.0062

TABLE XX
TRANSFORMER-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	170.9431	29221.5586	158.1415	-1.5834
Transformer(H=7)	170.6942	29136.5195	157.8691	-1.5707
Transformer(H=15)	170.1303	28944.3027	157.2814	-1.5470

TABLE XXI
TRANSFORMER-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	38.0804	1450.1200	34.1524	-0.0643
Transformer(H=7)	37.1832	1382.5886	33.3782	-0.0527
Transformer(H=15)	35.5135	1261.2057	32.0990	-0.0372

- **GRU**: Results for the GRU model.

TABLE XXII
GRU-1 (SMALL) TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	6.9945	48.9231	5.1213	0.9641
GRU(H=7)	17.1612	294.5072	13.9444	0.7758
GRU(H=15)	23.7065	561.9976	20.2164	0.5378

TABLE XXIII
GRU-1 (SMALL) TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	11.4523	131.1558	10.5839	0.9882
GRU(H=7)	51.1145	2612.6914	47.1789	0.7650
GRU(H=15)	92.2646	8512.7568	84.4537	0.2363

TABLE XXIV
GRU-2 (MEDIUM) TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	7.3000	53.2900	5.4166	0.9609
GRU(H=7)	17.0275	289.9360	13.7601	0.7792
GRU(H=15)	22.9595	527.1409	18.9595	0.5665

TABLE XXV
GRU-2 (MEDIUM) TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	10.5340	110.9660	9.6390	0.9901
GRU(H=7)	39.7209	1577.7537	36.7709	0.8594
GRU(H=15)	68.9906	4759.7065	63.4982	0.5770

TABLE XXVI
GRU-3 (LARGE) TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	8.3254	69.3125	6.4484	0.9491
GRU(H=7)	20.6921	428.1618	17.2917	0.6740
GRU(H=15)	30.5525	933.4529	26.1869	0.2324

TABLE XXVII
GRU-3 (LARGE) TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	32.5054	1056.6034	29.5108	0.9066
GRU(H=7)	93.9593	8828.3457	85.1432	0.2211
GRU(H=15)	134.4817	18085.3184	122.6083	-0.5914

- **Combined LSTM + GRU**: Results for the combined LSTM + GRU model.

TABLE XXVIII

COMBINED MODEL (LSTM+GRU) – SMALL: TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	9.8924	97.8598	7.9267	0.9282
LSTM+GRU(H=7)	21.2950	453.4767	17.5032	0.6547
LSTM+GRU(H=15)	31.3167	980.7339	26.1501	0.1935

TABLE XXIX

COMBINED MODEL (LSTM+GRU) – SMALL: TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	20.4317	417.4543	19.0524	0.9624
LSTM+GRU(H=7)	56.8397	3230.7556	52.4585	0.7094
LSTM+GRU(H=15)	99.8937	9978.7432	91.8942	0.1048

TABLE XXX

COMBINED MODEL (LSTM+GRU) – MEDIUM: TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	16.8905	285.2881	14.0157	0.7906
LSTM+GRU(H=7)	25.3874	644.5191	21.3244	0.5092
LSTM+GRU(H=15)	33.7993	1142.3899	28.6078	0.0605

TABLE XXXI

COMBINED MODEL (LSTM+GRU) – MEDIUM: TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	37.2281	1385.9325	34.6505	0.8751
LSTM+GRU(H=7)	58.1993	3387.1597	54.3907	0.6954
LSTM+GRU(H=15)	96.6440	9340.0537	90.0691	0.1621

TABLE XXXII

COMBINED MODEL (LSTM+GRU) – LARGE: TEST METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	37.3744	1396.8484	33.5533	-0.0252
LSTM+GRU(H=7)	36.4463	1328.3339	32.7445	-0.0114
LSTM+GRU(H=15)	34.8475	1214.3459	31.5199	0.0014

TABLE XXXIII

COMBINED MODEL (LSTM+GRU) – LARGE: TRAIN METRICS BY FORECAST HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	167.7451	28138.4199	154.7744	-1.4877
LSTM+GRU(H=7)	167.1399	27935.7422	154.1745	-1.4648
LSTM+GRU(H=15)	166.4760	27714.2441	153.5045	-1.4387

2) Multivariate Time-Series Forecasting: This section presents the key results for multivariate time-series forecasting using the same set of models.

- **LSTM**: Results for the multivariate LSTM model.

TABLE XXXIV

LSTM-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	10.0370	100.7411	8.3976	0.9909
LSTM(H=7)	14.1022	198.8712	12.1739	0.9821
LSTM(H=15)	13.2991	176.8671	11.3595	0.9841

TABLE XXXV

LSTM-1 (SMALL) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	7.5003	56.2551	5.5569	0.9587
LSTM(H=7)	8.8088	77.5943	6.5753	0.9409
LSTM(H=15)	8.4127	70.7742	6.2665	0.9418

TABLE XXXVI

LSTM-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	13.5444	183.4521	7.9137	0.9836
LSTM(H=7)	16.3880	268.5661	9.8617	0.9761
LSTM(H=15)	17.2119	296.2512	10.6690	0.9737

TABLE XXXVII

LSTM-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	12.9570	167.8826	10.5382	0.8768
LSTM(H=7)	15.2579	232.8045	12.4535	0.8227
LSTM(H=15)	15.6266	244.1894	12.7762	0.7992

TABLE XXXVIII

LSTM-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	74.2758	5516.9004	68.7854	0.5123
LSTM(H=7)	78.6276	6182.2988	72.7480	0.4545
LSTM(H=15)	85.4167	7296.0103	78.9118	0.3580

TABLE XXXIX

LSTM-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM(H=1)	22.7809	518.9694	19.4474	0.6191
LSTM(H=7)	23.5915	556.5605	20.0880	0.5762
LSTM(H=15)	24.3793	594.3495	20.8659	0.5112

- **Transformer**: Results for the multivariate Transformer model.

TABLE XL
TRANSFORMER-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	91.0024	8281.4414	86.0205	0.2538
Transformer(H=7)	97.5435	9514.7373	92.1294	0.1443
Transformer(H=15)	107.0506	11459.8311	101.0176	-0.0281

TABLE XLI
TRANSFORMER-1 (SMALL) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	61.7227	3809.6870	58.5992	-1.7961
Transformer(H=7)	65.9662	4351.5405	62.9058	-2.3134
Transformer(H=15)	72.3047	5227.9766	69.4631	-3.2993

TABLE XLII
TRANSFORMER-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	154.5453	23884.2617	143.9118	-1.1322
Transformer(H=7)	154.2494	23792.8691	143.6015	-1.1200
Transformer(H=15)	153.7909	23651.6445	143.1278	-1.1020

TABLE XLIII
TRANSFORMER-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	39.1745	1534.6393	34.2028	-0.1264
Transformer(H=7)	37.9190	1437.8483	33.1892	-0.0948
Transformer(H=15)	35.7663	1279.2310	31.6388	-0.0520

TABLE XLIV
TRANSFORMER-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	231.2129	53459.4102	214.9190	-3.7262
Transformer(H=7)	231.7780	53721.0352	215.3824	-3.7399
Transformer(H=15)	231.8493	53754.0859	215.3705	-3.7301

TABLE XLV
TRANSFORMER-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
Transformer(H=1)	46.6839	2179.3848	41.9397	-0.5996
Transformer(H=7)	46.3234	2145.8606	41.5193	-0.6339
Transformer(H=15)	45.3095	2052.9502	40.5851	-0.6883

- **GRU**: Results for the multivariate GRU model.

TABLE XLVI
GRU-MV-1 (SMALL) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	7.3669	54.2715	5.4517	0.9602
GRU(H=7)	16.6878	278.4828	13.2913	0.7880
GRU(H=15)	23.0331	530.5227	18.4586	0.5637

TABLE XLVII
GRU-MV-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	9.7639	95.3336	7.8696	0.9914
GRU(H=7)	16.1052	259.3788	11.1216	0.9767
GRU(H=15)	21.2739	452.5804	13.1856	0.9594

TABLE XLVIII
GRU-MV-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	7.2779	52.9685	5.2999	0.9611
GRU(H=7)	17.1691	294.7791	13.5292	0.7755
GRU(H=15)	23.9893	575.9196	19.1394	0.5264

TABLE XLIX
GRU-MV-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	9.9265	98.5346	8.8414	0.9912
GRU(H=7)	17.3801	302.0694	14.1956	0.9731
GRU(H=15)	23.0634	531.9199	17.0332	0.9527

TABLE L
GRU-MV-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	9.9379	98.7620	7.6766	0.9275
GRU(H=7)	16.7858	281.7632	13.4418	0.7855
GRU(H=15)	21.5104	462.6971	17.5309	0.6195

TABLE LI
GRU-MV-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
GRU(H=1)	19.4282	377.4534	17.3239	0.9666
GRU(H=7)	26.7059	713.2064	22.6479	0.9371
GRU(H=15)	29.7950	887.7421	24.2940	0.9219

- **Combined LSTM + GRU**: Results for the combined multivariate LSTM + GRU model.

TABLE LII
COMBINED-MV-1 (SMALL) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	7.8602	61.7828	5.8337	0.9547
LSTM+GRU(H=7)	17.0831	291.8325	13.5634	0.7778
LSTM+GRU(H=15)	24.0625	579.0057	19.0445	0.5238

TABLE LIII
COMBINED-MV-1 (SMALL) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	9.4442	89.1935	8.0371	0.9920
LSTM+GRU(H=7)	16.4850	271.7545	12.6596	0.9756
LSTM+GRU(H=15)	22.3044	497.4841	15.6595	0.9554

TABLE LIV
COMBINED-MV-2 (MEDIUM) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	9.9150	98.3081	7.7761	0.9278
LSTM+GRU(H=7)	16.7850	281.7347	13.4820	0.7855
LSTM+GRU(H=15)	21.8763	478.5728	17.7701	0.6064

TABLE LV
COMBINED-MV-2 (MEDIUM) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	13.8072	190.6395	9.8475	0.9830
LSTM+GRU(H=7)	19.4955	380.0743	13.2627	0.9661
LSTM+GRU(H=15)	22.9874	528.4219	14.5355	0.9530

TABLE LVI
COMBINED-MV-3 (LARGE) TEST METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	10.1295	102.6058	8.2022	0.9247
LSTM+GRU(H=7)	17.5101	306.6047	14.0696	0.7665
LSTM+GRU(H=15)	22.9732	527.7657	18.6077	0.5660

TABLE LVII
COMBINED-MV-3 (LARGE) TRAIN METRICS BY HORIZON

Model	RMSE	MSE	MAE	R^2
LSTM+GRU(H=1)	16.1287	260.1366	12.7257	0.9770
LSTM+GRU(H=7)	22.5764	509.6949	17.2358	0.9550
LSTM+GRU(H=15)	25.3818	644.2353	18.3042	0.9433

VIII. LIMITATIONS

In this project, we encountered several challenges in tuning the hyperparameters for each model, primarily due to the limited time.

Our focus was mainly on the OHLCAV data, and due to time constraints, we were unable to incorporate more advanced features into the models.

Initially, the performance of the models was suboptimal. This was primarily because a small portion of the Tesla dataset was used by mistake. Once we identified and corrected this issue, and utilized the full dataset from 2015 to 2024, the results improved significantly.

IX. VISUALIZATION AND ANALYSIS

A. Univariate time-series forecasting

1) LSTM: We have presented three different plots for the tree approaches in LSTM model.

Fig. 4. Training and testing error vs training time (LSTM) Approach 1 small

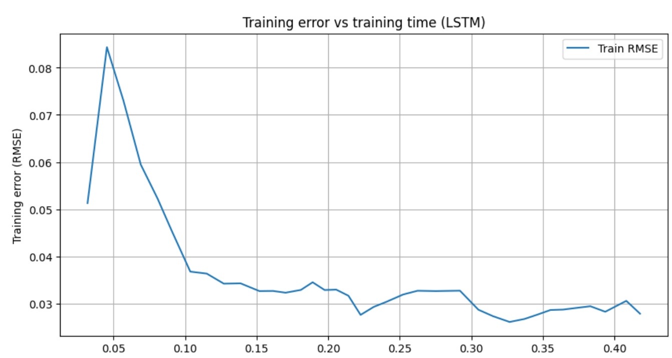

Fig. 5. Training error vs training time (LSTM) univariate Approach 1 small

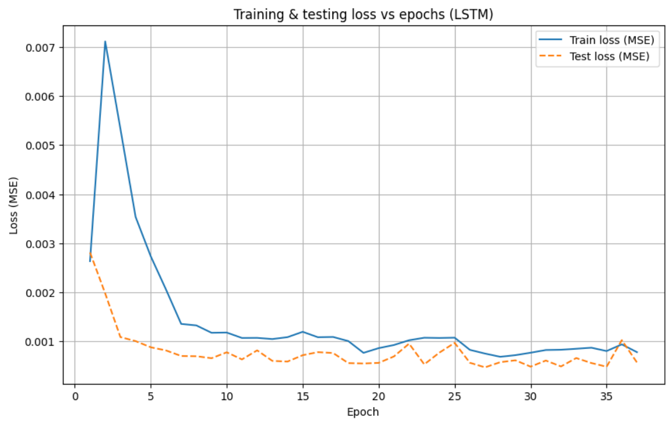

Fig. 6. Training and testing loss vs epochs (LSTM) univariate Approach 1 small

Fig. 7. Training and testing error vs training time (LSTM) univariate Approach 2 medium

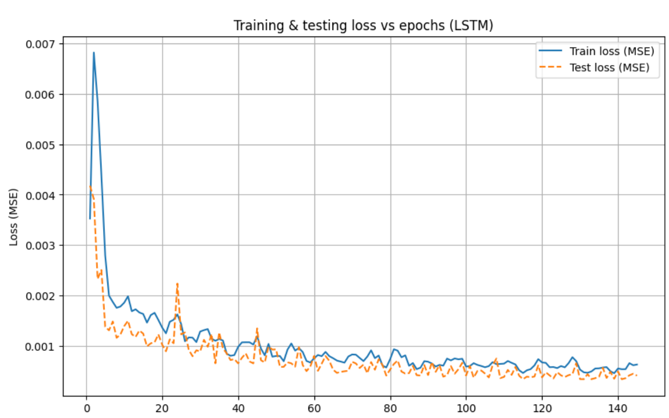

Fig. 8. Training and testing loss vs epochs (LSTM) univariate Approach 2 medium

Fig. 9. Training error vs training tim (LSTM) univariate Approach 2 medium

Fig. 10. Training and testing error vs training time Approach 3 large (LSTM)

Fig. 11. Training and testing loss vs epochs (LSTM) Approach 3 large

Fig. 12. Training error vs training time (LSTM) Approach 3 large

2) Transformer: In this section we will present the plots obtained from transformer model for each of the 3 approaches (small,medium,large)

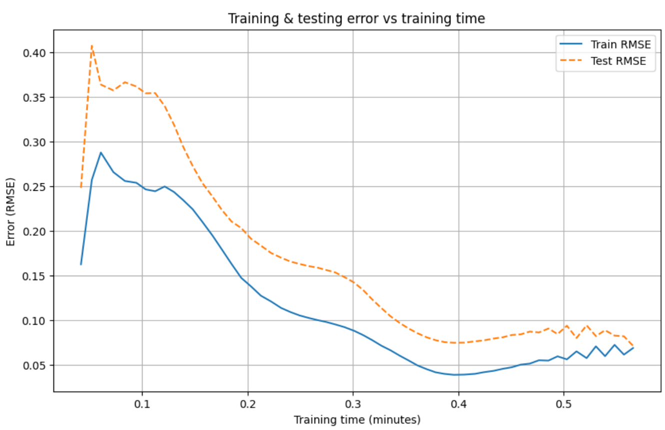

Fig. 13. Training and testing error vs training time transformer Approach 1 small

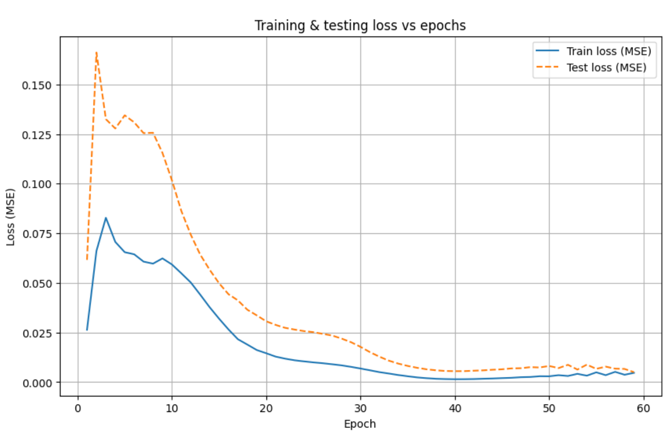

Fig. 14. Training and testing loss vs epochs transformer Approach 1 small

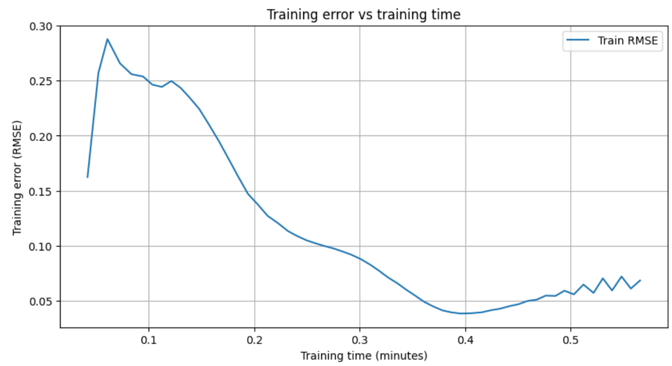

Fig. 15. Training error vs training time transformer Approach 1 small

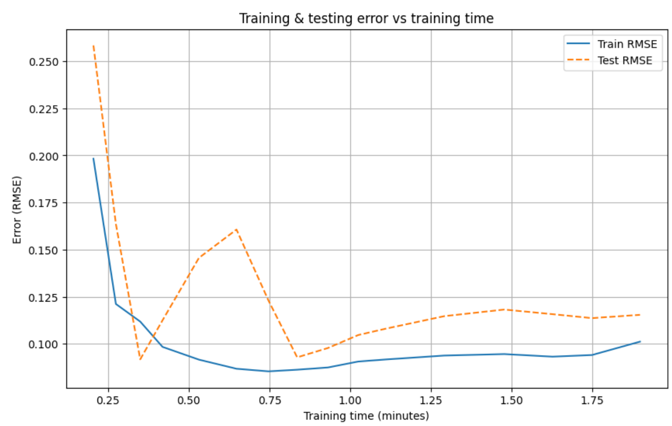

Fig. 16. Training and testing error vs training time transformer Approach 2 medium

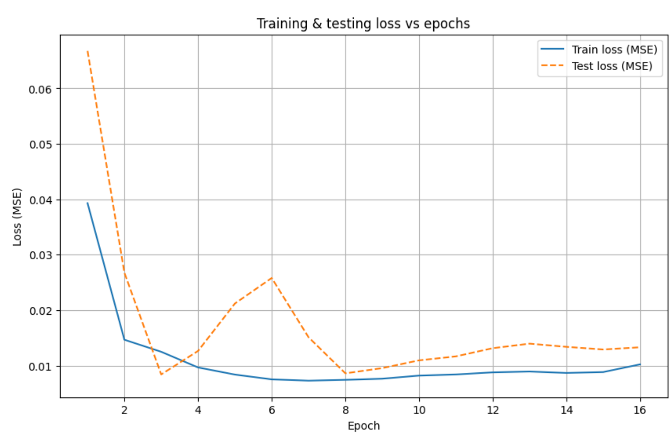

Fig. 17. Training and testing loss vs epochs transformer Approach 2 medium

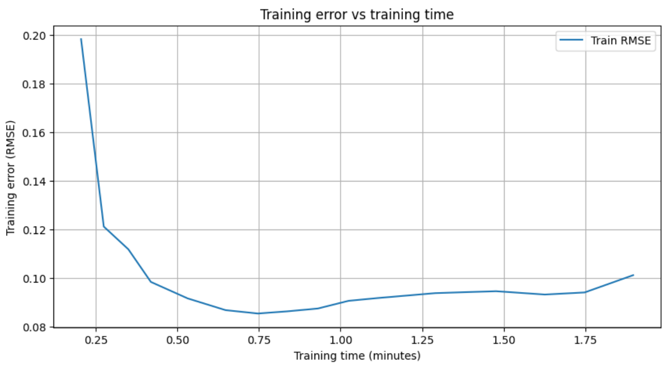

Fig. 18. Training error vs training time transformer Approach 2 medium

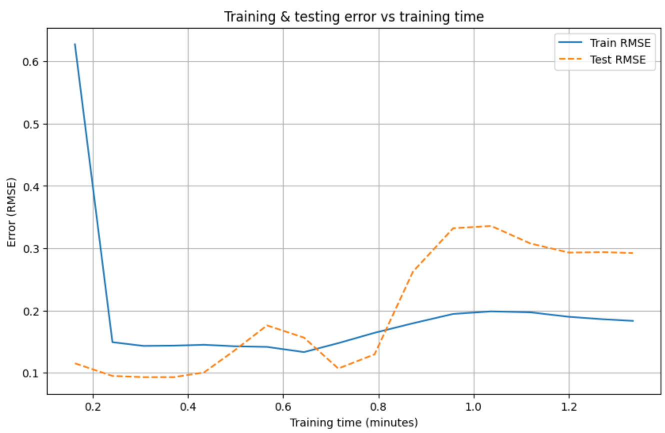

Fig. 19. Training and testing error vs training time transformer Approach 3 large

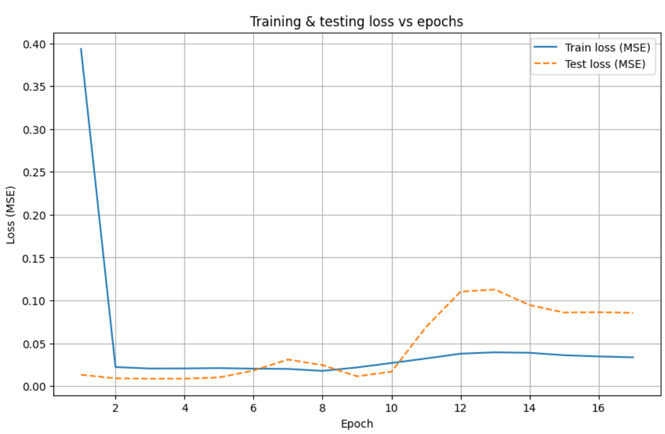

Fig. 20. Training and testing loss vs epochs transformer Approach 3 large

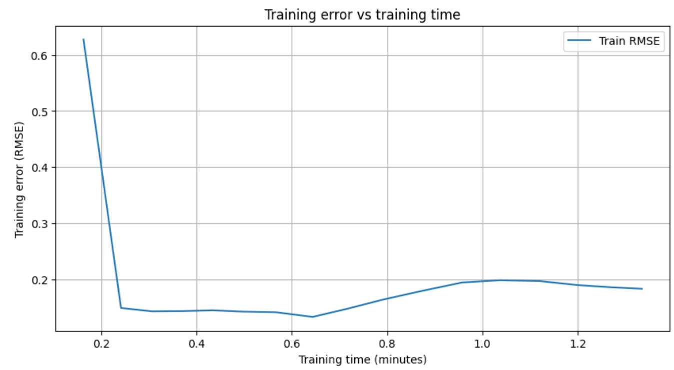

Fig. 21. Training error vs training time transformer Approach 3 large

3) GRU: In this section we will present the plots obtained from GRU model for each of the 3 approaches (small,medium,large)

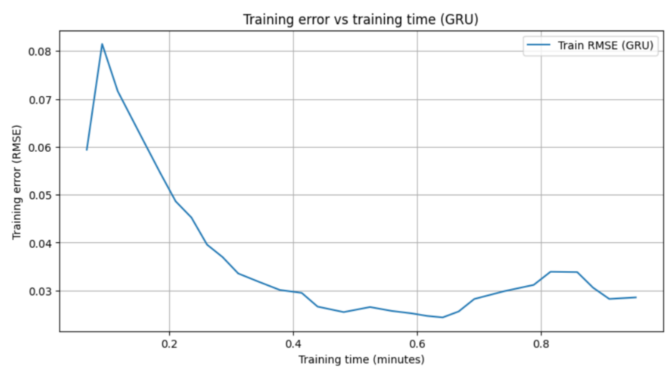

Fig. 22. Training error vs training time GRU Approach 1 small

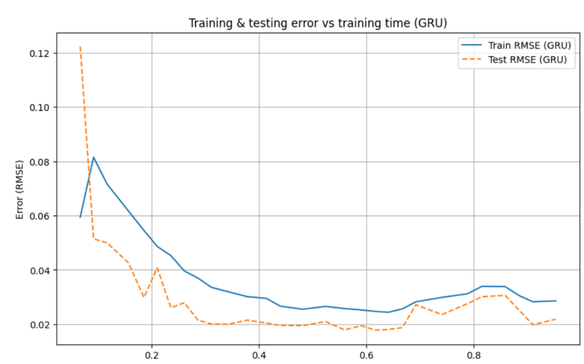

Fig. 23. Training and testing error vs training time (GRU) Approach 1 small

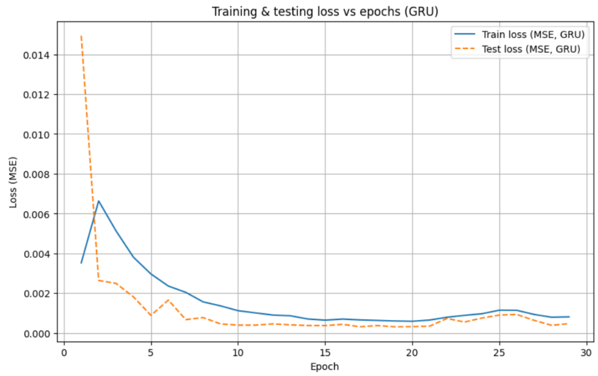

Fig. 24. Training and testing loss vs epochs (GRU) Approach 1 small

Fig. 25. Training error vs training time (GRU) Approach 2 Medium

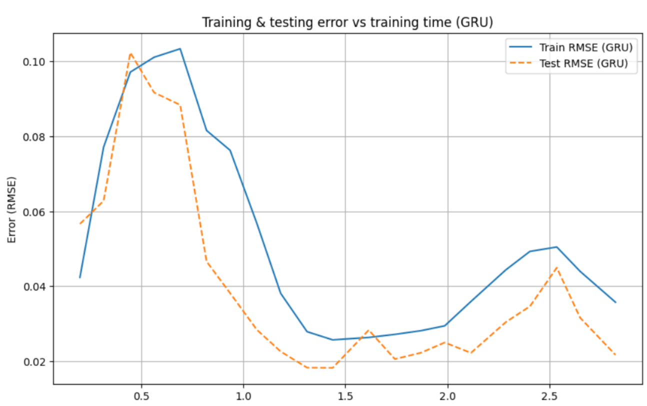

Fig. 26. Training and testing error vs training time (GRU) Approach 2 Medium

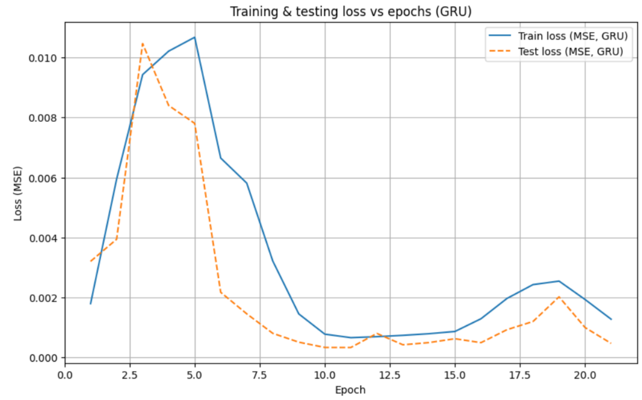

Fig. 27. Training and testing loss vs epochs (GRU) Approach 2 Medium

Fig. 28. Training error vs training time (GRU) Approach 3 Large

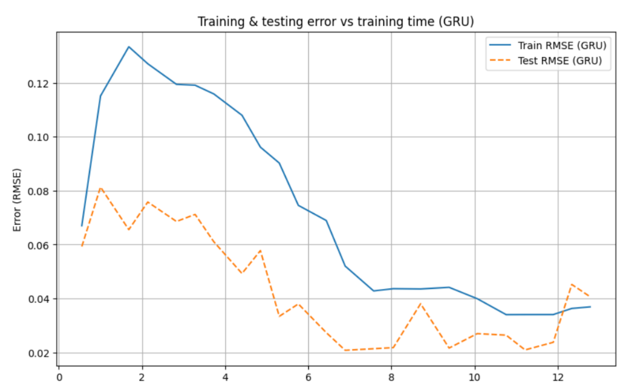

Fig. 29. Training and testing error vs training time (GRU) Approach 3 Large

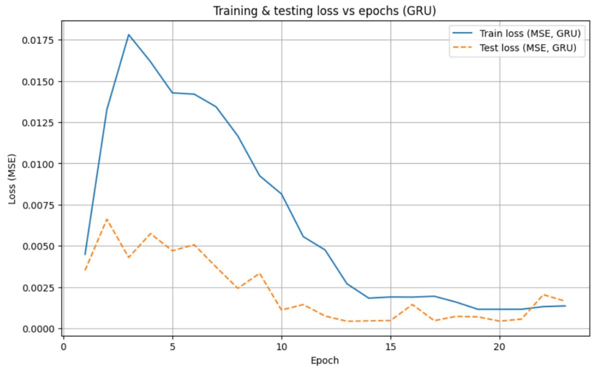

Fig. 30. Training and testing loss vs epochs (GRU) Approach 3 Large

4) Combined LSTM + GRU: In this section we will present the plots obtained from combined model for each of the 3 approaches (small,medium,large)

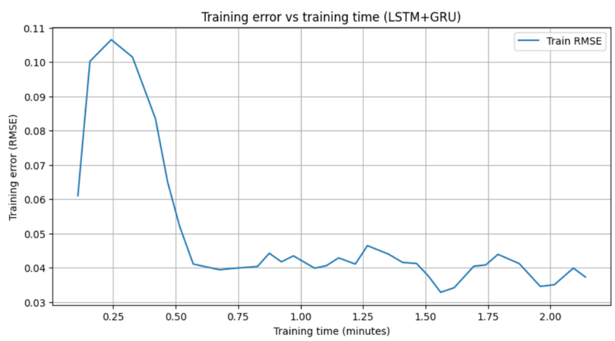

Fig. 31. Training error vs training time (combined) Approach 1 small

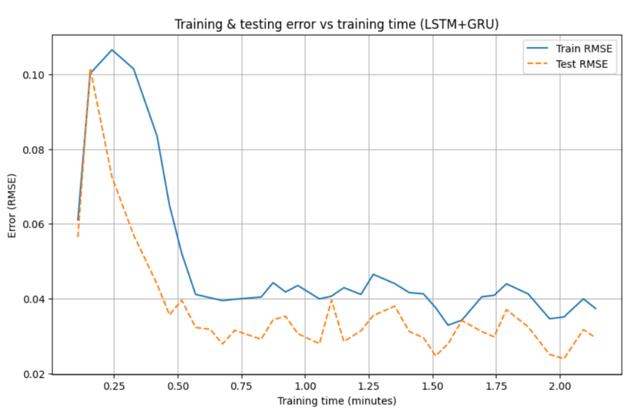

Fig. 32. Training & testing error vs training time (combined) Approach 1 small

Fig. 33. Training & testing loss vs epochs (combined) Approach 1 small

Fig. 34. Training error vs training time (combined) Approach 2 medium

Fig. 35. Training & testing error vs training time (combined) Approach 2 medium

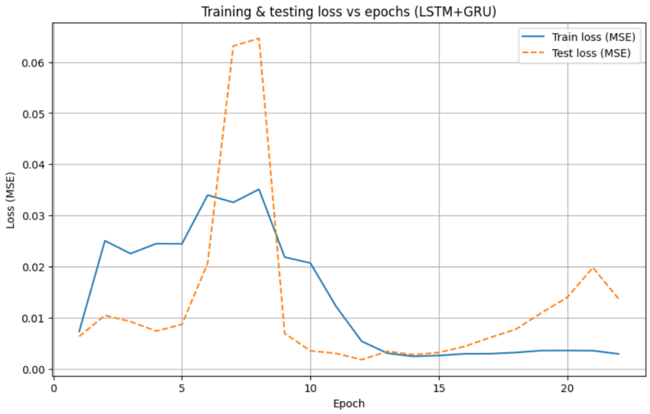

Fig. 36. Training & testing loss vs epochs (combined) Approach 2 medium

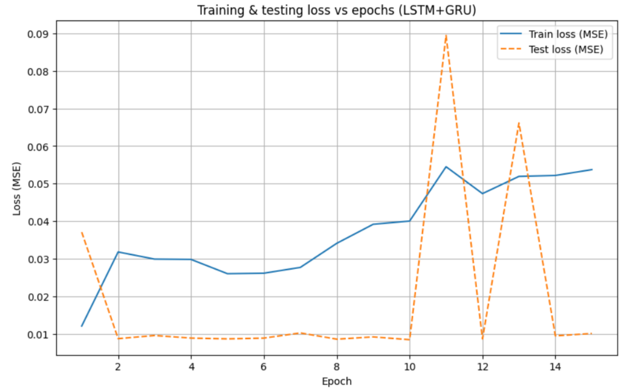

Fig. 39. Training & testing loss vs epochs (combined) Approach 3 large

B. Multivariate time-series forecasting

1) LSTM: Those are the plots for LSTM multivariate.

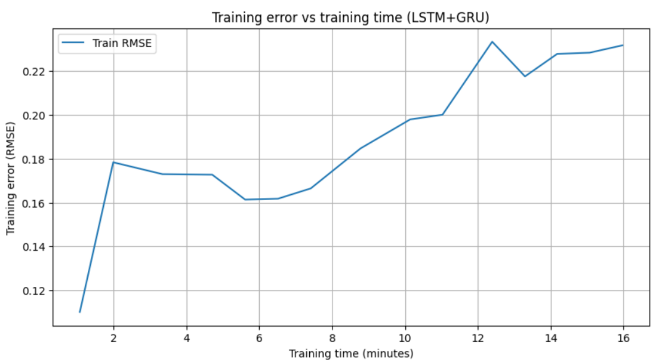

Fig. 37. Training error vs training time (combined) Approach 3 large

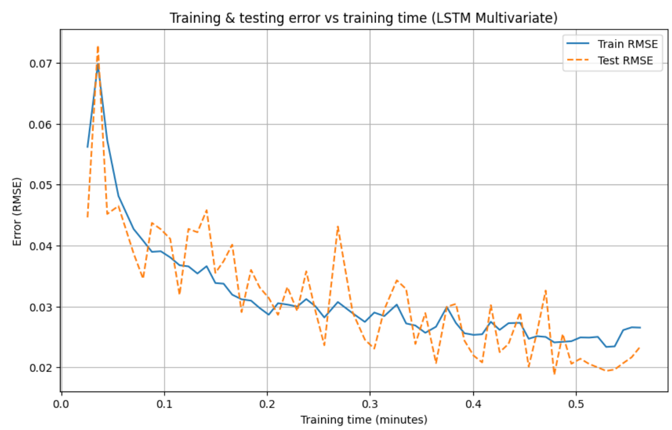

Fig. 40. Training and testing error vs training time (LSTM Multivariate) approach 1

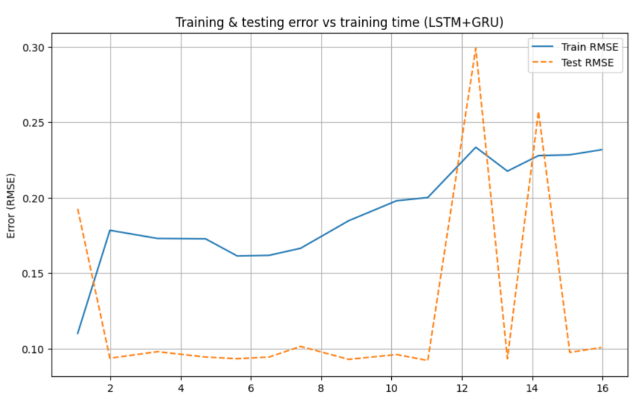

Fig. 38. Training & testing error vs training time (combined) Approach 3 large

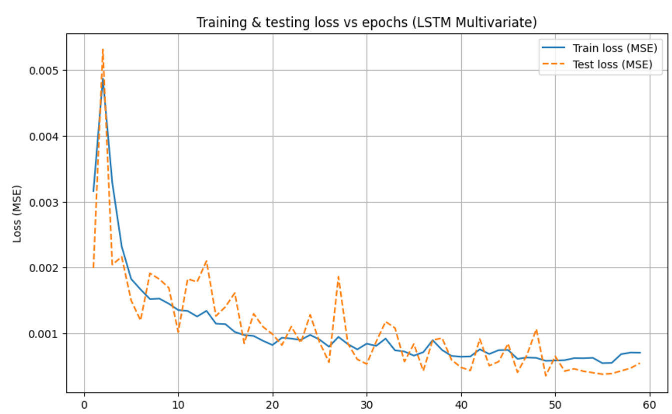

Fig. 41. Training and testing error vs training time (LSTM Multivariate) approach 1

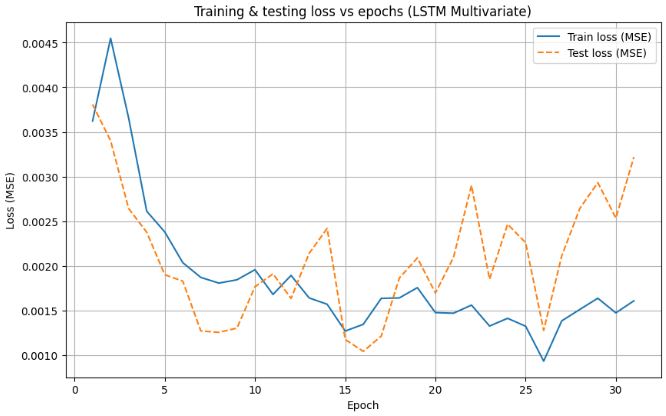

Fig. 42. Training and testing loss vs epochs (LSTM Multivariate) approach 2

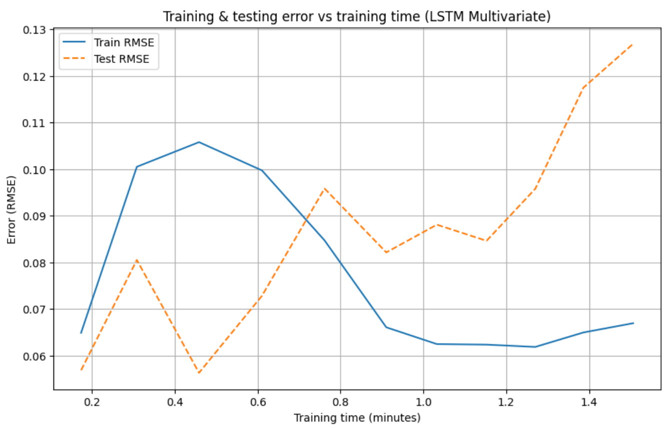

Fig. 43. Training and testing error vs training time (LSTM Multivariate) approach 3

Fig. 44. Training and testing loss vs epochs (LSTM Multivariate) approach 3

2) Transformer: Those plots are for transformer with multivariate.

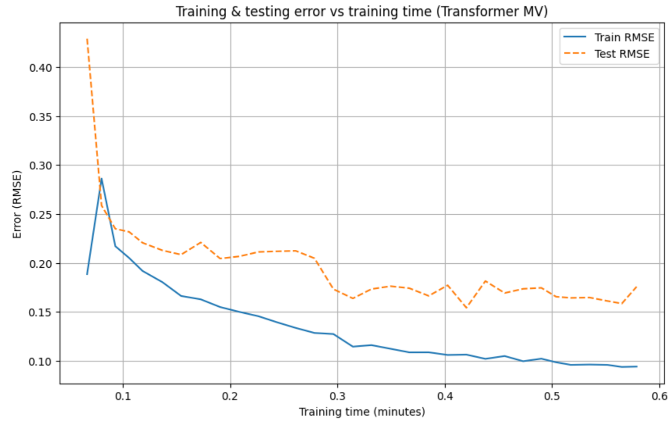

Fig. 45. Training and testing error vs training time (Transformer MV) first approach

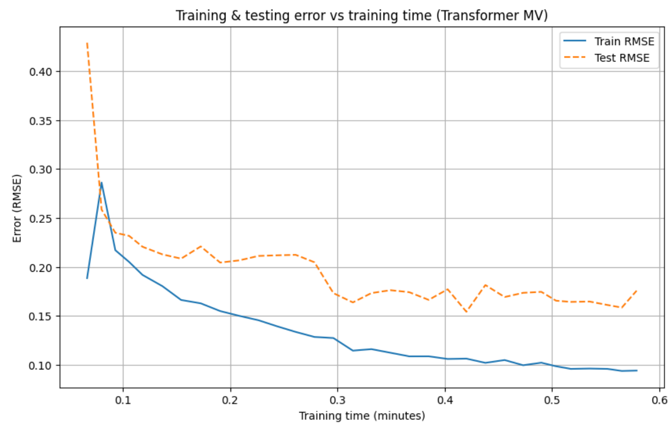

Fig. 46. Training and testing error vs training time (Transformer MV) first approach

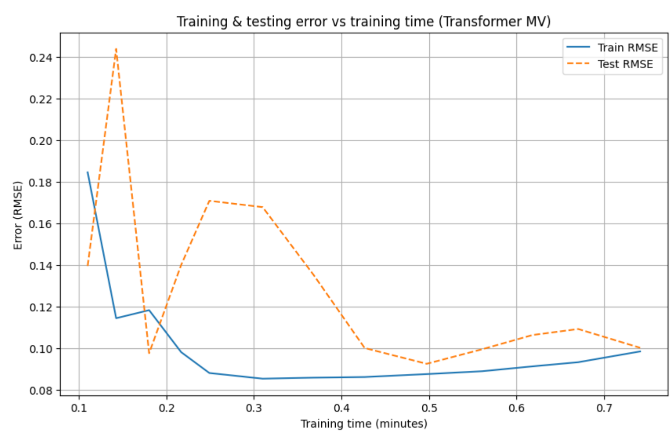

Fig. 47. Training and testing error vs training time (Transformer MV) second approach

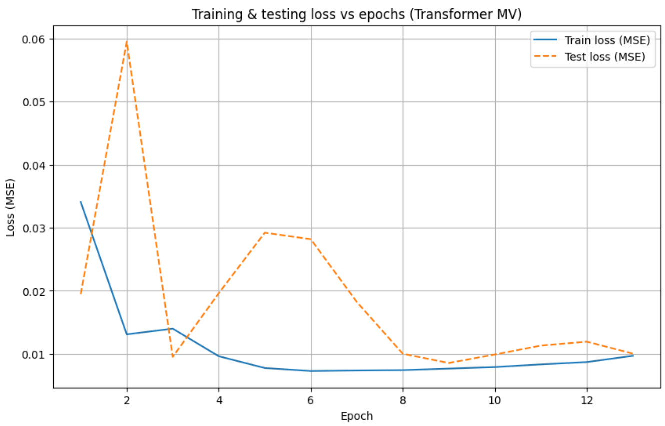

Fig. 48. Training and testing loss vs epochs (Transformer MV) second approach

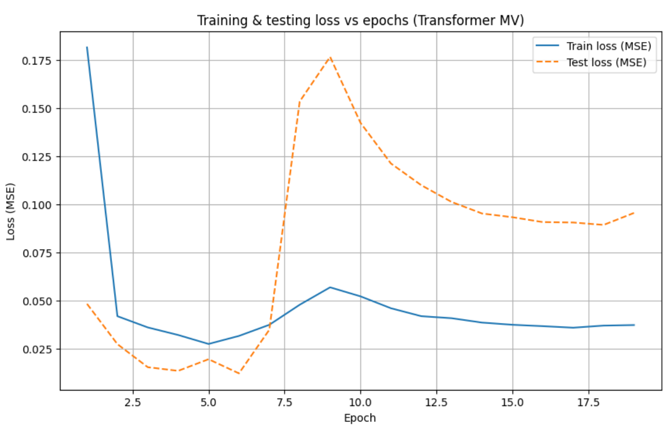

Fig. 49. Training and testing loss vs epochs (Transformer MV) third approach

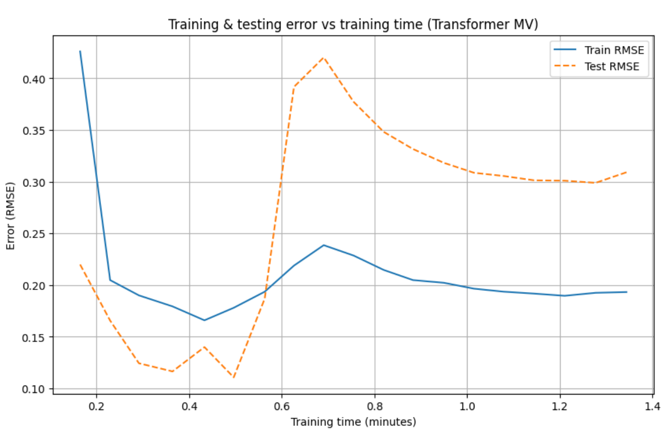

Fig. 50. Training and testing error vs training time (Transformer MV) third approach

3) *GRU:* Those are the plots for GRU multivariate.

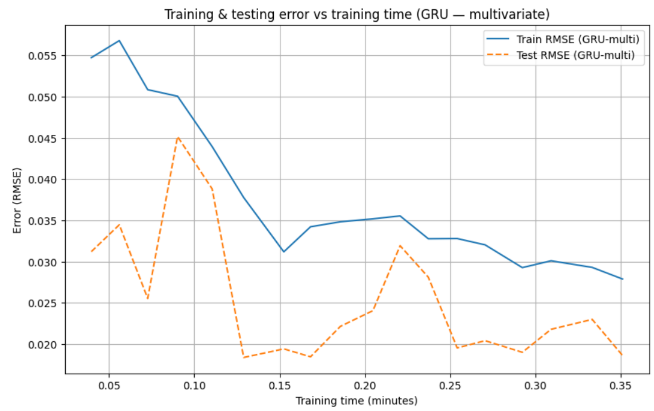

Fig. 51. Training & testing error vs training time (GRU MV) first approach

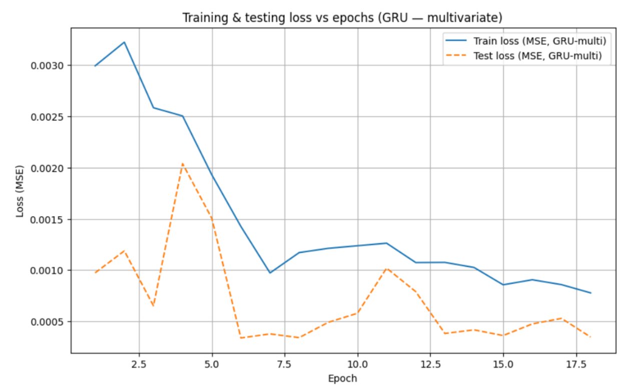

Fig. 52. Training and testing loss vs epochs (GRU MV) first approach

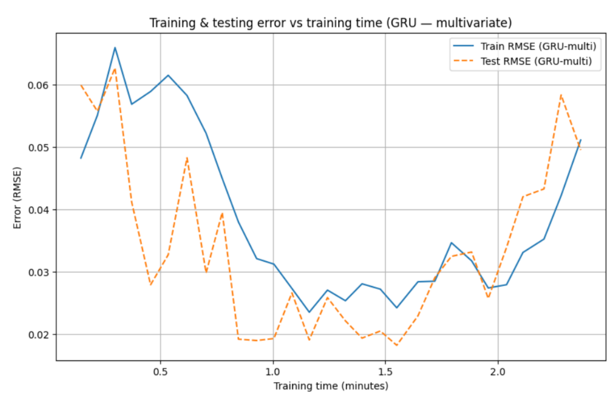

Fig. 53. Training & testing error vs training time (GRU MV) second approach

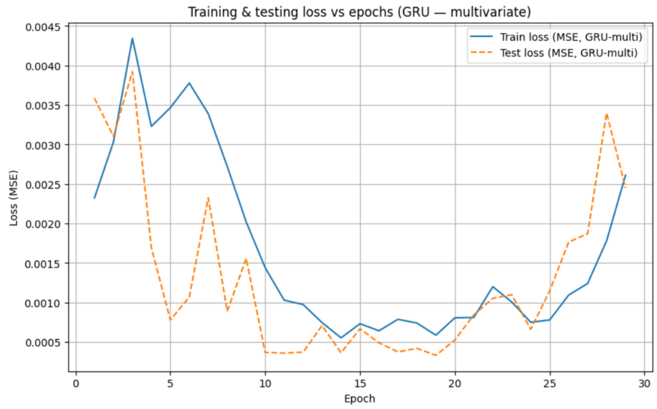

Fig. 54. Training and testing loss vs epochs (GRU MV) second approach

Fig. 55. Training & testing error vs training time (GRU MV) third approach

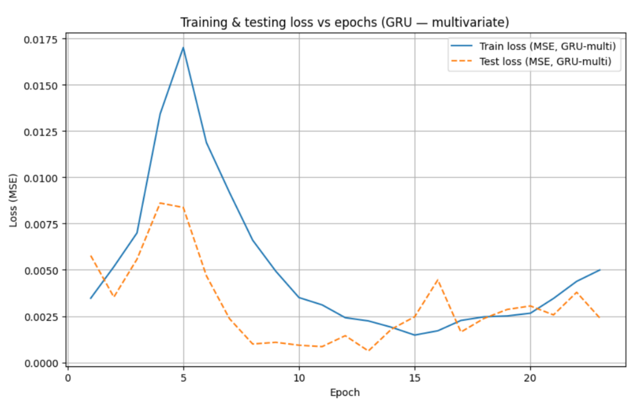

Fig. 56. Training and testing loss vs epochs (GRU MV) third approach

4) Combined LSTM + GRU: Those are the plots for combined multivariate.

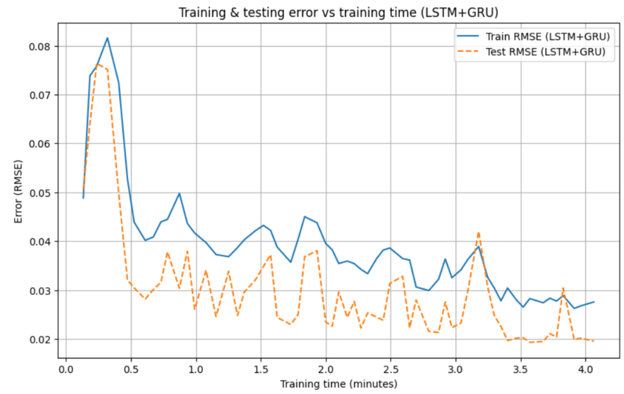

Fig. 57. Training & testing error vs training time (combined MV) first approach

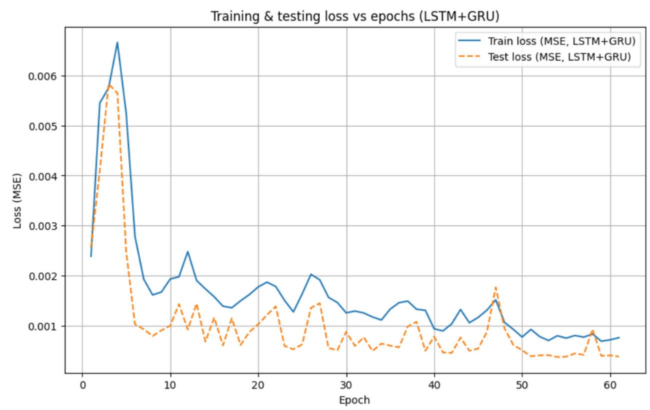

Fig. 58. Training and testing loss vs epochs (combined MV) first approach

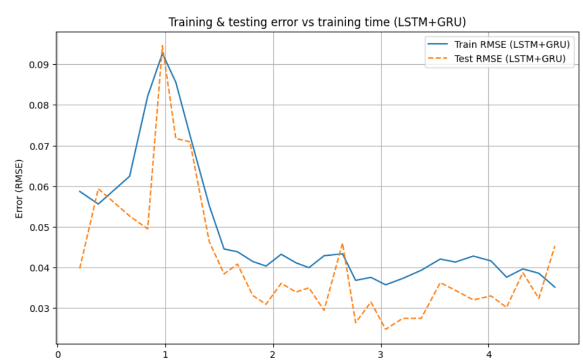

Fig. 59. Training & testing error vs training time (combined MV) second approach

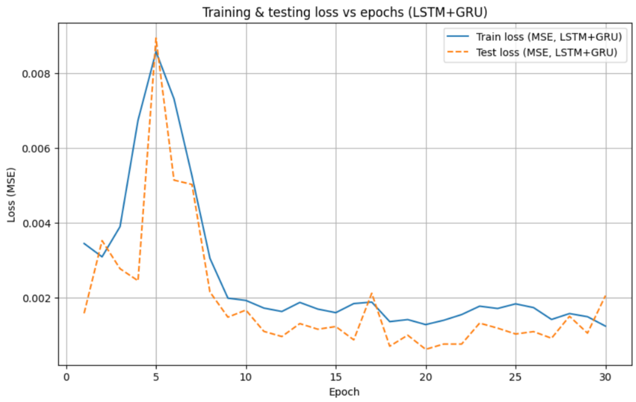

Fig. 60. Training and testing loss vs epochs (combined MV) second approach

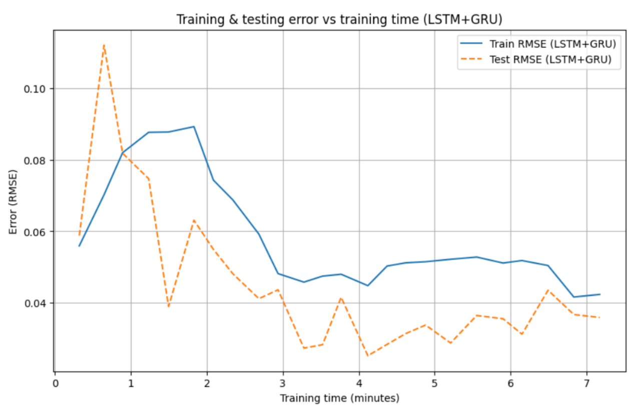

Fig. 61. Training & testing error vs training time (combined MV) third approach

Fig. 62. Training and testing loss vs epochs (combined MV) third approach

X. Conclusion

In this work, we compared LSTM, GRU, LSTM+GRU, and Transformer architectures for forecasting Tesla stock prices using univariate and multivariate time series from 2015–2024. Across all configurations and horizons, the best overall performance was obtained by the GRU second approach (medium) for one-step-ahead forecasting ($H=1$), achieving RMSE $=$ 7.3000, MSE $= 53.2900$, MAE $= 5.4166$, and $R^2 = 0.9609$ on the test set. Small multivariate LSTM models further showed that incorporate OHLCAV features can stabilize multi-step forecasts, while combined (LSTM + GRU) and Transformer models did not outperform simpler GRU and LSTM.

References

[1] S. Prabhu and J. Kundargi, "Stock market prediction using univariate LSTM model," *International Journal of Innovative Science, Engineering & Technology*, vol. 8, no. 8, pp. 361–371, 2021, iSSN 2348-7968.

[2] P. Srivastava and P. K. Mishra, "Stock market prediction using RNN LSTM," in *2021 2nd Global Conference for Advancement in Technology (GCAT)*. Bangalore, India: IEEE, 2021, dOI available via IEEE Xplore.

[3] Y. Chi, "Predictive analysis of Tesla's stock closing prices utilizing LSTM and GRU deep learning models," in *Proceedings of the 1st International Conference on Data Science and Engineering (ICDSE 2024)*. SCITEPRESS, 2024, pp. 422–428.

[4] E. T. Sivadasan, N. M. Sundaram, and R. Santhosh, "Stock market forecasting using deep learning with long short-term memory and gated recurrent unit," *Soft Computing*, vol. 28, no. 6, pp. 3267–3282, 2024.