Stock Forecasting using GRU based Models

Darsh Rathi (2022A7PS0187P), Harsh Shah (2022A7PS0169P) and Malay Mishra (2022A7PS0116P)

April 22, 2025

Abstract :

Stock Forecasting is a standard time series analysis problem. It has applications in the real world for earning money too. We begin with analysis of a existing research paper on CNN-BiLSTM-ECA model, and replicating its results. We then move on to testing a CNN-GRU model, a purely GRU based model, a Transformer model and a TCN model. The GRU model by far outperforms all other models with an MAE of 0.012 and the CNN-GRU model is a close second with an MAE of 0.023. An ensemble model of GRU and CNN-GRU model also performs quite well with an MAE of 0.014.

1 Introduction

The prediction of the stock market is crucial for financial strategy. Stock market prices are highly nonlinear, noisy, and nonstationary.Traditional statistical models struggle with nonlinearities. Deep learning can capture temporal patterns and complex dependencies. Initially this project was based on the ARIMA model, which is a comparatively simpler model for Time Setries Analysis. But current research focuses on Time series analysis using Deep learning and RNNs. The Base paper we focused on is **Stock Price Forecast Based on CNN-BiLSTM-ECA Model** by Yu Chen et al. This paper proposed a BiLSTM model, integrated with a CNN layer to capture short term dependencies and an Efficient Attention layer to emphasize dependencies in the data. This model with 10 timesteps has an MAE of 0.028, and in our testing had an MAE of 0.044.

2 Methods

Our dataset has been obtained from Kaggle, it consists of the Stock market data of the NIFTY50 companies from JAN 2001 to DEC 2021. Dataset Link Here

2.1 Base Paper

A hybrid model is introduced with three main components and 4 layers:
CNN Layer: Extracts spatial and short-term patterns from the input time window. Uses 1D convolution over a fixed-length window of input prices.Helps reduce noise and enhances local trend capture.
BiLSTM Layer:Captures long-term dependencies in both forward and backward directions.Overcomes the limitation of unidirectional LSTM which misses future context.Outputs hidden states that encapsulate richer temporal information.
ECA Module (Efficient Channel Attention):Applies a lightweight channel-wise attention mechanism.Dynamically assigns importance weights to feature channels.Helps the model focus on more informative features.
Output Layer:Fully connected (dense) layer maps features to predicted stock price values.
Their Data is from the Shanghai Stock Exchange. This data is not available publicly.

2.2 Implementation of CNN-GRU

We implemented CNN-GRU, taking inspiration from their CNN-BiLSTM model. Similar to their approach, the CNN layer captures short term dependencies and the GRU layer focuses on extracting longer dependencies and important features. The model has a 1D Convolutional Layer(3x3 kernel, ReLU activation), a 64 unit GRU Layer, a 32 Unit GRU layer, a 5 parameter Efficient Channel Attention and Pooling and a Dense layer to finalise the results. similar to the base paper.

2.3 Implementation of GRU

We also implemented a GRU based model with no CNN layer, to compare results of the CNN based layer. It consist of 2 layers of GRUs(64 units and 32 units) and a Dropout layer with a Dense Layer at the end to compile the results.

3 Ablation

For the ablation study of our proposed model, we tried 7 variations on the current model, whose descriptions and results are as follows.

- Removing the second layer of the GRU model.
- Making the same model with LSTMs.
- Removing the Hidden Dense layer
- Changing the Activation to tanh
- Changing dropout rate to 0.5.
- Changing the optimiser to RMSProp
- Changing the optimiser to SGD

Model Results	MSE	MAE
Baseline GRU	0.0012	0.0214
no hidden dense	0.004	0.047
optimizer rmsprop	0.005	0.044
no second gru	0.006	0.043
lstm	0.007	0.059
optimizer sgd	0.009	0.055
tanh activation	0.011	0.052
dropout 05	0.015	0.092

4 Results

The average results of the models over all stocks of the NIFTY50 dataset are as follows.

Model Results	MSE	MAE	RMSE
GRU	0.0012	0.0214	0.0267
Ensemble	0.0014	0.0240	0.0317
CNN-GRU-ECA	0.0023	0.0317	0.0418
CNN-BiLSTM-ECA	0.048	0.0410	0.0521

5 Conclusion

This report proposes a GRU based model for Stock forecasting and shows that it works better over other Deep Learning based models.