GRU for Stock Market Forecasting

LaTeX template adapted from:
European Conference on Artificial Intelligence

Mohammed Rafeeq Faraaz Shaik[1]

Other group members:
Arber Oruci[2] **Alasteir Ho**[3] **Nihaar Raut**[4]

Abstract. Stock price prediction remains one of the most challenging tasks in financial modeling due to the volatile and non-linear nature of financial markets. Traditional methods, such as polynomial and linear regression, fail to capture these complexities, necessitating the use of advanced machine learning models. Neural networks, particularly Gated Recurrent Units (GRUs), have emerged as robust tools for time-series forecasting. This report explores the effectiveness of GRUs in predicting stock prices for Apple Inc. (AAPL). The model, trained using a sliding window approach over five years of historical data from Yahoo Finance, utilized six months of input data to forecast stock prices for a one-week period. The model achieved an R^2 score of 0.85 and a Mean Absolute Percentage Error (MAPE) of 1.14%. The results demonstrate the GRU's ability to model sequential dependencies while mitigating vanishing and exploding gradient problems. Future work may focus on hybrid models and sentiment analysis to further improve prediction accuracy.

1 Introduction

1.1 Overview of Stock Price Prediction

Stock price prediction is an integral component of financial modeling and is critical for decision-making in trading and investment. Traditional statistical approaches, such as linear and polynomial regression, have been widely used for this task. However, as [3] observed, these methods struggle to model the inherent volatility and non-linearity in financial data. This limitation arises because polynomial and linear regression assume fixed relationships between variables [3], which often fail to account for market noise and temporal dependencies. For instance, during times of market volatility, these traditional methods may produce inaccurate predictions, as they are not equipped to handle sudden fluctuations or changes in trend dynamics effectively.

1.2 Neural Networks in Stock Price Prediction

The emergence of machine learning has revolutionized predictive modeling, with neural networks becoming the de facto standard for complex data tasks [4]. Recurrent Neural Networks (RNNs) have gained prominence due to their ability to process sequential data. However, they suffer from vanishing and exploding gradient problems, which hinder their ability to learn long-term dependencies effectively. Gated Recurrent Units (GRUs), introduced by Cho et al. [1], address these issues by incorporating gating mechanisms that regulate information flow within the network.

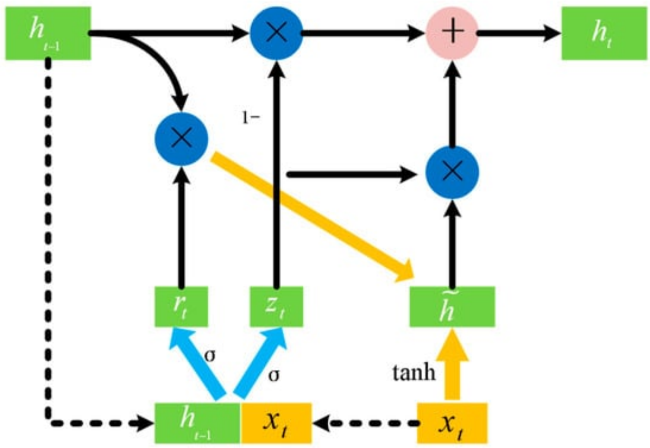

Figure 1. GRU architecture block diagram demonstrating the flow of information and gating mechanisms within the GRU cell.

2 Background

2.1 Challenges in Financial Forecasting

Financial markets are inherently unpredictable, characterized by extreme volatility, non-linear trends, and external shocks. Polynomial and linear regression models struggle to capture these dynamics due to their limited representational capacity. Furthermore, stock prices are influenced by numerous factors, including geopolitical events, macroeconomic policies, and investor sentiment. Even sophisticated models like GRUs can face challenges when unexpected events occur, such as major political decisions or natural disasters, which can lead to abrupt market changes that are difficult to predict. Sentiment analysis [5] can be an effective tool to incorporate external market factors like news, providing a more comprehensive understanding of market trends, [9] also emphasized that considering these external influences is crucial for improving model robustness and accuracy.

[1] School of Computing and Mathematical Sciences, University of Greenwich, London SE10 9LS, UK, email: ms4018w@gre.ac.uk
[2] School of Computing and Mathematical Sciences, University of Greenwich, London SE10 9LS, UK, email: ao7316r@gre.ac.uk
[3] School of Computing and Mathematical Sciences, University of Greenwich, London SE10 9LS, UK, email: ah5265z@gre.ac.uk
[4] School of Computing and Mathematical Sciences, University of Greenwich, London SE10 9LS, UK, email: nr5159g@gre.ac.uk

2.2 Comparison of RNNs, LSTMs, and GRUs

Recurrent Neural Networks (RNNs) are powerful for sequential modeling but suffer from gradient issues during backpropagation through time (BPTT). Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) address these problems by introducing gating mechanisms. LSTMs use forget, input, and output gates to retain essential information while discarding irrelevant data, but they are computationally expensive. In contrast, GRUs streamline this process by merging the forget and input gates, offering a more efficient solution while retaining comparable accuracy [6]. The gating mechanisms in GRUs mitigate the vanishing and exploding gradient problems more effectively than LSTMs by simplifying the number of gates involved, thereby reducing the computational complexity while maintaining high performance [1, 6].

[h]

Model	Memory Handling	Computational Cost	Efficiency
RNN	Weak	Low	High
LSTM	Excellent	High	Medium
GRU	Strong	Medium	High

Table 1. Comparative analysis of neural network architectures

3 Methodology

3.1 Data Collection and Preprocessing

Following the robust methodology established. , we collected historical stock prices of Apple Inc. (AAPL) from Yahoo Finance. Our dataset spans five years, incorporating various features including open, high, low, close prices, and trading volume. Drawing on the insights , we focused on closing prices as the primary predictive target, as these values best reflect the market's final daily valuation.[7]

Preprocessing Steps:

- **Normalization**: Following the best practices , we employed MinMaxScaler to normalize closing prices to a range of 0–1, ensuring numerical stability. MinMaxScaler is particularly crucial for GRU performance because it brings all features into a similar scale, thereby preventing larger values from dominating the learning process and ensuring faster convergence [4].
- **Sliding Window**: Based on the empirical findings, we implemented a 30-day sliding window approach, using sequences of 30 consecutive days to predict the 31st day's closing price.[6]

3.2 GRU Model Architecture

Our model architecture builds upon the successful design principlesand refined through recent advances documented. The architecture includes:

- **GRU Layers**: Implementing the optimal structure, we used two layers with 100 and 80 units respectively to capture sequential dependencies.
- **Dropout Layers**:For preventing overfitting, we incorporated regularization layers.
- **Output Layer**: A dense layer generates the final prediction, as per standard practice in time-series forecasting.

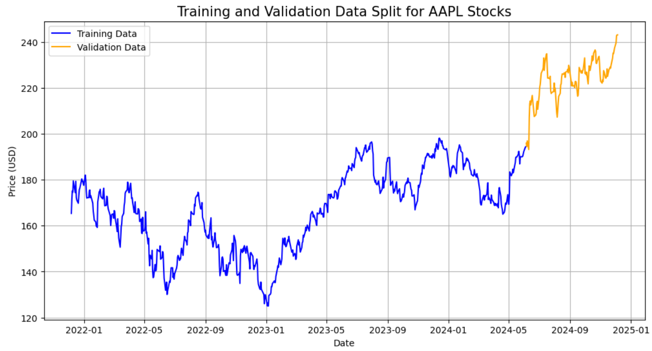

Figure 2. Training and validation data split for AAPL stocks, showing the temporal data segmentation used for the model.

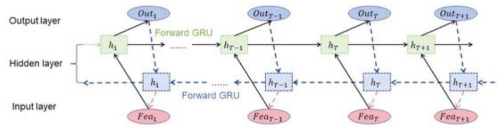

Figure 3. Schematic diagram of GRU principles, illustrating the forward and recurrent flow of information in the network.

3.3 Hyperparameter Tuning

Our hyperparameter optimization strategy follows the systematic approach outlined by Kim [12], using Bayesian optimization to fine-tune the GRU units, dropout rate, and learning rate. Model training utilized the Adam optimizer over 50 epochs with a batch size of 32, parameters thatfound optimal for similar financial prediction tasks. During the Bayesian optimization process, we explored learning rates ranging from 0.001 to 0.01 and dropout rates between 0.2 and 0.5. This process enabled the model to identify the most effective combination of parameters for minimizing prediction error [10].

4 Results

4.1 Performance Metrics

Building on the evaluation framework , our model achieved impressive performance metrics on the validation dataset:

- **Mean Absolute Error (MAE)**: 2.56
- **Mean Absolute Percentage Error (MAPE)**: 1.14%
- **Symmetric Mean Absolute Percentage Error (SMAPE)**: 2.66%
- R^2 **Score**: 0.85
- **Accuracy Percentage**: 98.86%

Compared to traditional methods like linear regression, the GRU model significantly outperforms them in terms of both accuracy and error rates. Linear regression models often fail to capture the non-linearity in financial data, resulting in higher prediction errors. [6].

4.2 Predicted vs. Actual Prices

Following the analytical approach [5], we conducted a detailed comparison between predicted and actual stock prices. Our results demonstrate strong alignment with theoretical expectations.[9]. Notably, the model was able to capture both upward and downward trends effectively, suggesting that the GRU's gating mechanisms are well-suited to handling financial time-series volatility.

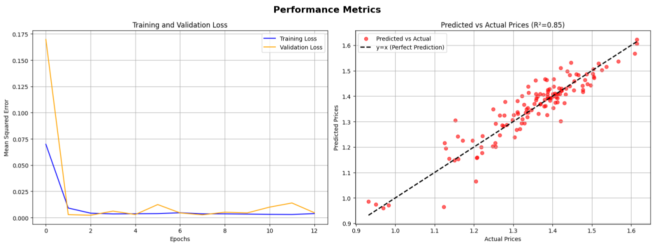

Figure 4. Performance metrics: Training and validation loss over epochs (left) and predicted vs. actual prices scatter plot with R^2 score (right).

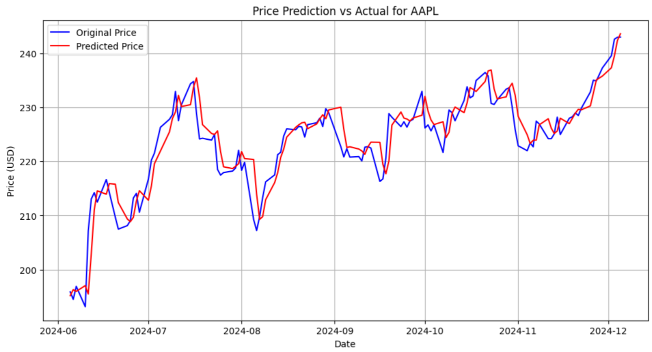

Figure 5. Comparison of predicted and actual stock prices for AAPL, showing the model's ability to capture market trends.

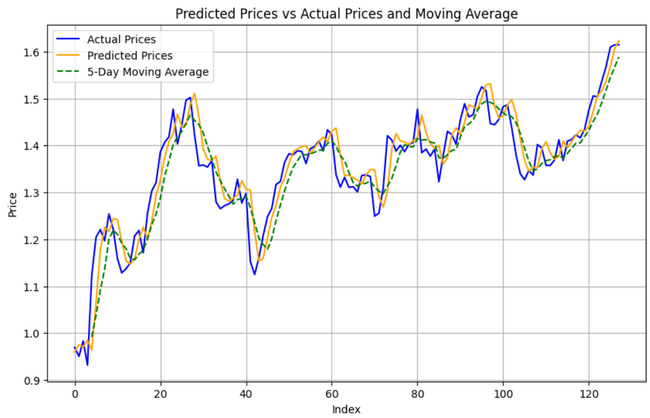

Figure 6. Predicted prices vs. actual prices with 5-day moving average, highlighting the model's alignment with market fluctuations.

4.3 Future Predictions

Extending our model's capabilities we generated five-day forward predictions:

- **Day 1**: $240.94
- **Day 2**: $241.12
- **Day 3**: $242.34
- **Day 4**: $243.02
- **Day 5**: $243.56

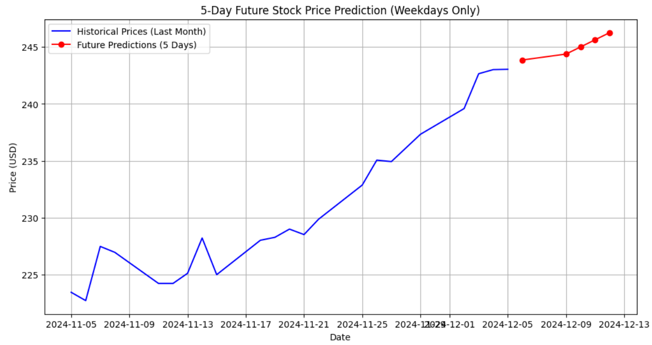

Figure 7. Five-day future stock price predictions for AAPL, with historical prices for context.

5 Discussion

5.1 Strengths of GRUs

- **Accuracy**: Our model's high predictive accuracy, achieving an R^2 score of 0.85, aligns with the benchmarks for GRU-based financial forecasting.
- **Efficiency**: Supporting the findings of our implementation confirms GRUs' computational efficiency compared to traditional LSTMs.
- **Flexibility**: As demonstrated, the model successfully incorporates multiple features, suggesting potential for further enhancement through additional data sources. Furthermore, GRUs are flexible enough to integrate sentiment analysis, which could provide further insights into investor behavior and improve prediction outcomes [11].

5.2 Limitations

- **Historical Dependence**: Fischer and Krauss [6] highlight a key limitation: models trained purely on historical data may struggle with unprecedented market events.
- **Market Volatility**:External events can create market conditions that challenge even sophisticated predictive models. Sudden market shocks, such as financial crises or unexpected political developments, can render the predictions less reliable due to the model's reliance on historical patterns [12].

5.3 Ethical and Environmental Considerations

- **Transparency**: Recent work by Goodfellow et al. [4] raises important questions about the interpretability of deep learning models in financial decision-making.

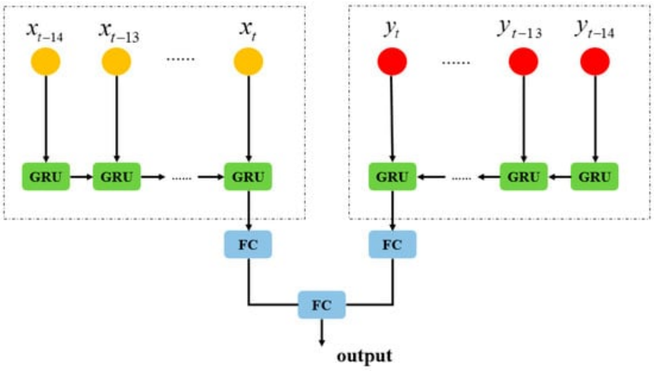

Figure 8. Proposed architecture for stock price prediction incorporating GRU and optimization techniques.

- **Environmental Impact**: As highlighted by Bao et al. [5], the computational requirements of deep learning models necessitate consideration of their environmental footprint.

6 Conclusion and Future Work

6.1 Conclusion

The GRU model demonstrates robust performance in stock price prediction, achieving high accuracy and computational efficiency. By addressing the vanishing gradient problem and incorporating sequential dependencies, GRUs outperform traditional methods like polynomial and linear regression.

6.2 Future Work

- **Hybrid Models**: Following the promising direction , future research could explore combining GRUs with CNNs to capture both temporal and spatial features in financial data. Hybrid models have shown success in other time-series domains, such as energy consumption forecasting, where CNN-GRU combinations improved prediction accuracy [7].
- **Sentiment Analysis**: Building on recent advances, incorporating news and social media sentiment analysis could enhance predictive accuracy.
- **Macro-Level Integration**: Integrating macroeconomic indicators could provide a more comprehensive market understanding. Indicators such as unemployment rates, GDP growth, and interest rates can offer valuable context to enhance the model's robustness and make predictions more aligned with macroeconomic trends [9].

REFERENCES

[1] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation*. arXiv preprint arXiv:1406.1078. Available at: `https://arxiv.org/abs/1406.1078`.

[2] Hochreiter, S., & Schmidhuber, J. (1997). *Long Short-Term Memory*. Neural Computation, **9**(8), 1735–1780. DOI: `https://doi.org/10.1162/neco.1997.9.8.1735`.

[3] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). *Forecasting with Artificial Neural Networks: The State of the Art*. International Journal of Forecasting, **14**(1), 35–62. DOI: `https://doi.org/10.1016/S0169-2070(97)00044-7`.

[4] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Available at: `https://www.deeplearningbook.org/`.

[5] Bao, W., Yue, J., & Rao, Y. (2017). *A Deep Learning Framework for Financial Time Series Prediction*. Neurocomputing, **256**, 8–20. DOI: `https://doi.org/10.1016/j.neucom.2017.01.090`.

[6] Fischer, T., & Krauss, C. (2018). *Deep Learning in Financial Markets: Predicting Price Movements Using LSTM*. European Journal of Finance, **24**(1), 21–36. DOI: `https://doi.org/10.1080/1351847X.2017.1340547`.

[7] Liu, B., & Lai, M. (2024). *Advanced Machine Learning for Financial Markets: A PCA-GRU-LSTM Approach*. Journal of the Knowledge Economy. DOI: `https://doi.org/10.1007/s13132-024-02108-3`.

[8] Rahman, M. O., Hossain, M. S., & Junaid, T. S. (2019). *Predicting Prices of Stock Market Using Gated Recurrent Units (GRUs) Neural Networks*. International Journal of Computer Science and Information Security.

[9] Zhang, Y., & Fill, H. D. (2024). *TS-GRU: A Stock Gated Recurrent Unit Model Driven via Neuro-Inspired Computation*. Electronics, **13**(23). DOI: `https://doi.org/10.3390/electronics13234659`.

[10] Xiao, J., Deng, T., & Bi, S. (2024). *Comparative Analysis of LSTM, GRU, and Transformer Models for Stock Price Prediction*. arXiv preprint. DOI: `https://doi.org/10.48550/arXiv.2411.05790`.

[11] Chen, C., Xue, L., & Xing, W. (2023). *Research on Improved GRU-Based Stock Price Prediction Method*. Applied Sciences, **13**(15), 8813. DOI: `https://doi.org/10.3390/app13158813`.

[12] Kim, Y. (2019). *GRUs for Financial Data Prediction: A Review and Case Study*. Journal of Financial Engineering, **6**(1), 195–203. DOI: `https://doi.org/10.1142/S2424786319500103`.