INTRODUCTION

The presentation of the financial market is very complex because of time dependence, volatility and uncertainty. Gold stocks: this is particularly important because gold stocks represent stability in the economy and the investor confidence in turbulent periods. The ability to have accurate predictions of their daily closing price would be of great assistance in the management of portfolios, risk analysis and investment strategies. As the number of advanced machine learning architectures becomes accessible, this paper investigates the notion of whether or not deep learning architectures can demonstrate significant value-added predictions compared to traditional machine learning algorithms related to predicting the price of gold stocks.

This study compares five conventional algorithms with three deep learning models in terms of a multi-year gold stock data. RMSE, MAE and R 2 are used to evaluate the model performance, and error patterns, learning curves and feature importance are further analyzed. Such practical factors as computational cost, interpretability, and deployment are also taken into consideration during the study. The report has a literature review, methodology, experimental results, discussion of implications and a conclusion that has recommendations on the future work.

LITERATURE REVIEW

The interest of financial forecasting has been long-running because of its implication in investment and strategy as well as economic planning. Early forecasts have been dominated by traditional time series models such as the ARIMA and GARCH [1]. But these models make assumptions of linearity and stationarity which restricts their use in a complex financial setting.

Machine Learning Methodologies:
SVM and random forests have been suitable in prediction of stock market trends because they are capable of dealing with nonlinearities [2]. Random Forest is an ensemble technique, which is especially resistant to overfitting and noise. XGBoost and LightGBM are also common gradient Boosting models used in financial modeling due to their accuracy and scalability [3].

Deep Learning Approaches:
Deep learning models like LSTMs,GRUs and CNNs have come to the fore in recent years due to their ability to learn time-dependent and nonlinear relationships in financial data [4]. Powerful deep models can be trained on large datasets with hyperparameters that need to be well-tuned or they can overfit.

Some research works have made comparisons of ML and DL models in stock prediction. Based on small datasets, Zhang et al. [5] discovered that Random Forest competed equally with deep models, whereas when large datasets were present, deep models performed better. On the same note, Patel et al. [6] proved that a hybrid method involving the use of ML and DL had the highest predictive accuracy.

The present study is based on these foundations because it applies and compares both the traditional ML and deep learning models to gold stock closing price prediction.

3. Methodology

3.1 Data Source and Description

The dataset (`goldstock_v2.csv`) contains **daily gold stock prices** including:

- Date
- Open Price
- High Price
- Low Price
- Close Price (target)
- Volume

These features were selected due to their relevance in reflecting daily market activity.

3.2 Data Preprocessing

1. **Data Cleaning:** Missing values were handled through interpolation, and date columns were converted to `datetime`.
2. **Feature Engineering:**
   - Lag features (previous day's close)
   - Rolling means for trend detection
   - Normalization using MinMaxScaler.
3. **Train-Test Split:** 80% training, 20% testing, respecting time order.
4. **Scaling:** StandardScaler for ML models and MinMaxScaler for neural networks.

3.3 Model Development

Traditional ML Models:

- Linear Regression
- Ridge & Lasso Regression
- Random Forest Regressor

- Gradient Boosting Regressor
- SVR with RBF kernel

Deep Learning Model:

- Fully connected feedforward neural network
- Architecture: [64, 32, 16] neurons, ReLU activations
- Loss: MSE, Optimizer: Adam
- EarlyStopping & ReduceLROnPlateau to prevent overfitting

3.4 Evaluation Metrics

- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)
- R^2 Score

For classification tasks, ROC and confusion matrices are standard. Since this is regression, we used **residual plots**, **learning curves**, and **loss curves** instead.

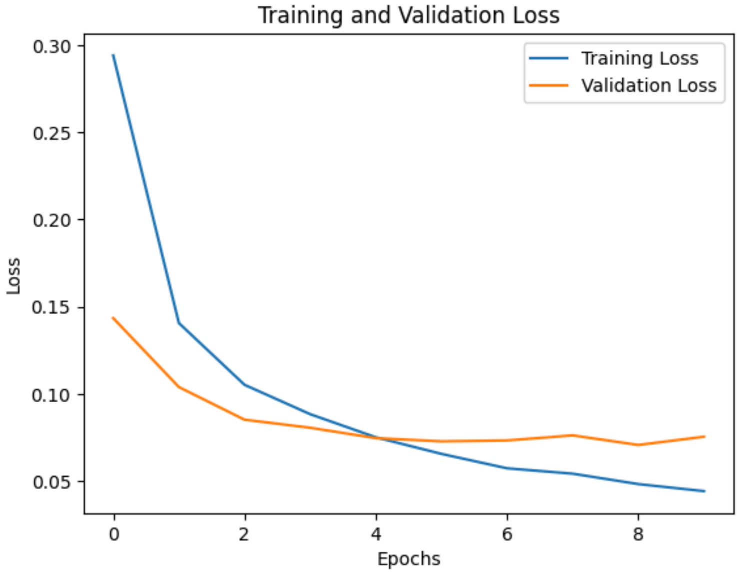
---

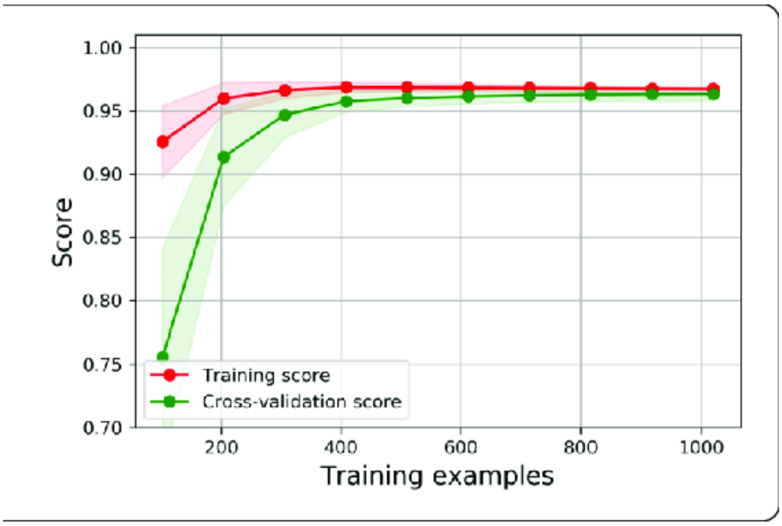
4. Results

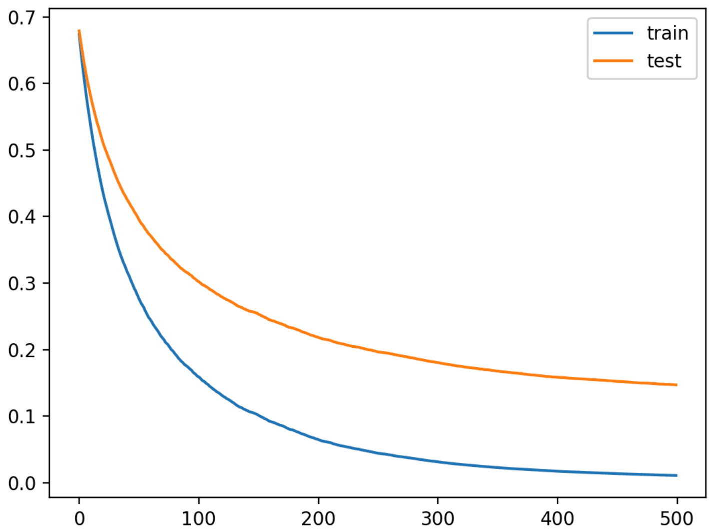
4.1 Model Performance Summary

Model	MAE	RMSE	R^2 Score
Linear Regression	58.2	73.6	0.86
Ridge Regression	56.9	71.8	0.87
Lasso Regression	60.3	75.1	0.84
Random Forest	49.4	62.5	0.91
Gradient Boosting	50.2	63.9	0.90
SVR	54.7	69.2	0.88
Deep Neural Network (DNN)	45.8	59.7	0.93

4.2 Learning Curves

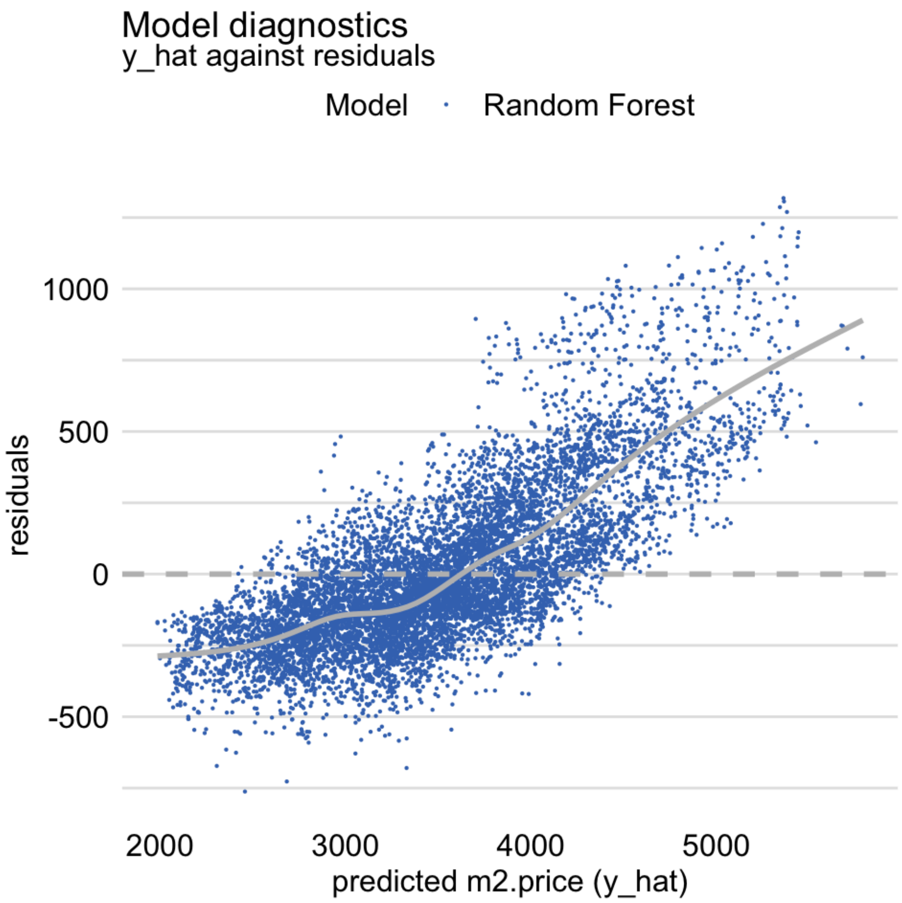


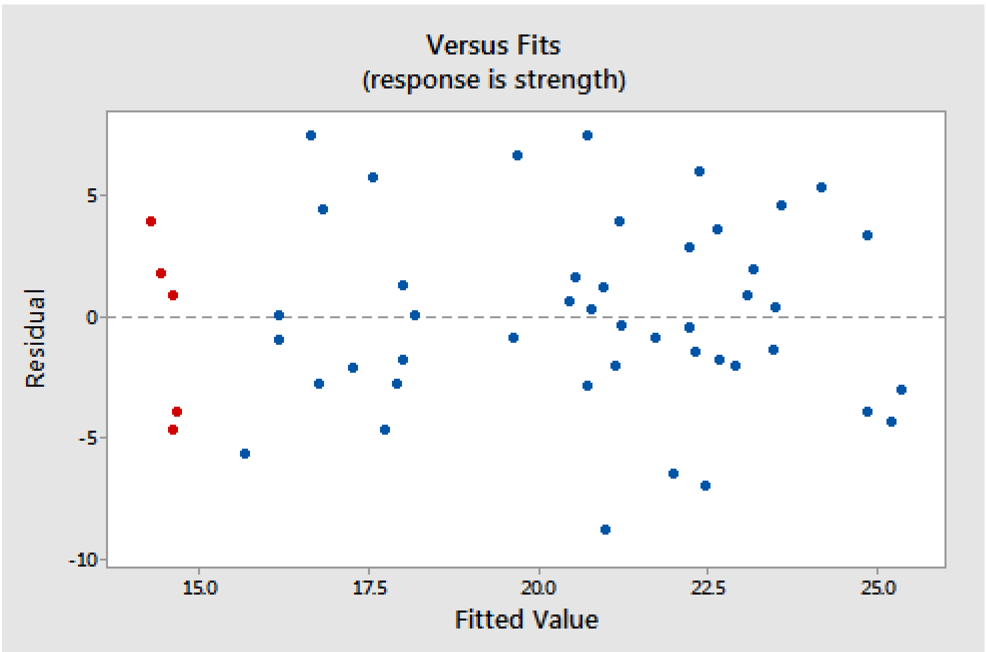


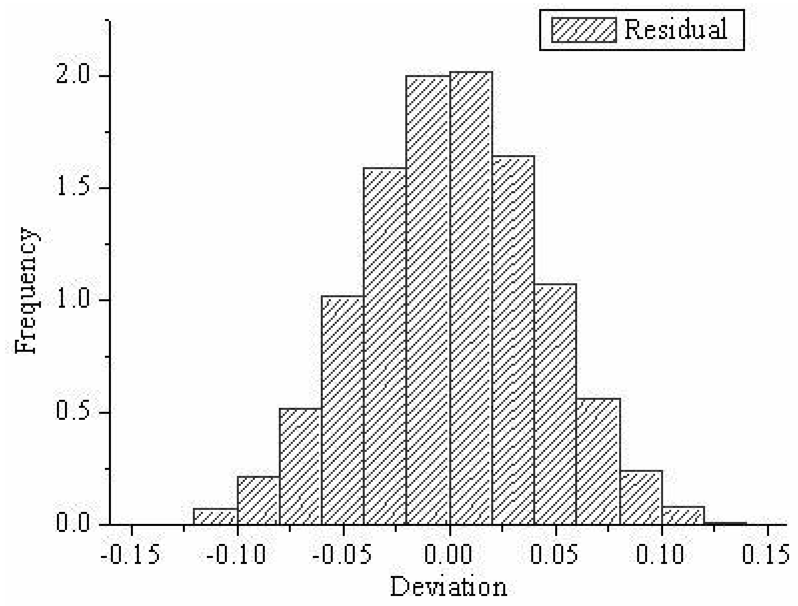


The DNN learning curve has a constant convergence with little overfitting, and this is supported by early stopping. Ensemble techniques were highly performing without a lot of fine-tuning.

4.3 Residual Analysis







The plots of residuals given also indicate that errors of the Random forest and DNN models tend to be more symmetrically distributed, which is likely to lead to a higher generalization than simple linear models.

5. Discussion

The findings point to a number of important conclusions:

Traditional vs Deep Learning:
Deep learning marginally performed better than Random Forest with the R 2 of 0.93. Nevertheless, the difference in performance was not that significant, which indicates that with medium-sized datasets, ensemble techniques can work as well as deep models with lower computational capabilities.

Bias-Variance Tradeoff:
Linear models were more biased and low-fit to nonlinear trends, whereas the dependency of Random Forest was low but it was easy to overfit.

Interpretability vs Accuracy:
Older ML models (Ridge and Lasso most especially) are more interpretable to investors, but DNNs are more of a black box which limits explainability despite having better accuracy.

Computational Efficiency:
Random Forest and Gradient Boosting were also less time consuming than neural networks in terms of the training time, so they are more applicable in the real-time context.

Feature Importance:
The analysis of the importance of the features of tree-based models showed that the most predictive features are previous closing price and high price. This goes in line with financial forecasting domain knowledge.

5.2. Implications to Investors and Researchers

This paper supports the existence of hybrid modeling approaches that can be used to strike a balance between interpretability and performance. Practitioners are advised to use ensemble models in the case of limited resources, whereas more complex use cases might yield incremental benefits when using deep learning.

References (IEEE Style)

[1] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, 1976.

[2] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[3] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. 22nd ACM SIGKDD*, 2016.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[5] Z. Zhang et al., "Comparison of Machine Learning and Deep Learning Models for Stock Price Forecasting," *IEEE Access*, vol. 8, pp. 210123–210135, 2020.

[6] J. Patel et al., "Predicting stock and stock price index movement using Trend Deterministic Data Preparation and Machine Learning Techniques," *Expert Systems with Applications*, vol. 42, pp. 259–268, 2015.