Using Neural Network Architectures for Intraday Trading in the Gold Market

Srinivas Devarajula, Vitaliy Milke\textsuperscript{a} and Cristina Luca\textsuperscript{b}

Computing and Information Science, Anglia Ruskin University, East Road, Cambridge, U.K.

Keywords: Machine Learning in Finance, Intraday Trading, Neural Network Computing, Artificial Intelligence, Algorithmic Trading, Forex.

Abstract: Financial market forecasting is used to assess the future value of financial instruments in various exchange and over-the-counter markets. Investors have a high interest in the most accurate prediction of the financial instruments’ prices. Inaccurate forecasting might result in a significant financial loss in certain circumstances. This research aims to determine the most probabilistic deep learning model that can improve price forecasting in the financial markets. In this research, Convolutional Neural Networks and Long Short-Term Memory are used for the experiments to forecast the Gold price movements on the Forex market. The Gold(XAU/USD) dataset is used in this research to predict the prices for the next minute. The models proposed have been evaluated using Mean squared error, Mean absolute error, and Mean absolute percentage error metrics. The results show that the Convolutional Neural Network has performed better than the Long Short-Term Memory network and has the potential to predict the price for next minute with a low error rate.

1 INTRODUCTION

In financial markets, there are several ways to invest based on the length of time one investor (natural or legal person) holds financial instruments such as stocks, bonds, futures, options, commodities or Forex. In long-term investment, the investors hold the financial assets for months or years, whilst intraday trading refers to the investment done in financial instruments for minutes or hours per day. While both strategies have benefits and drawbacks, intraday trading is regarded as riskier due to the volatility of market changes (Bhat and Kamath, 2013), (Demirer et al., 2021).

Futures and Forex markets prediction is one of the most challenging problems due to the data volatility and high level of noise in the data. Adding unbalanced data makes predicting the accurate price movement for the next few minutes even harder. As the market behaviour is constantly changing, the patterns identified could be altered too. There are studies on time series prediction using the minute-to-minute dataset for currencies ((Evans, 2018), (Raimundo and Okamoto Jr, 2018), (Weeraddana et al., 2018), (Chen et al., 2019), (Rundo et al., 2019), (Islam and Hossain, 2020), (Liao et al., 2021)), Standard Poor’s 500 (S&P 500) index-based financial instruments, such as SPY Exchange-Traded Fund (ETF) and E-Mini futures ((Ferreira and Medeiros, 2021), (Kinyua et al., 2021)), and studies of commodities and precious metal markets using classical statistical methods ((Huang et al., 2021), (Semeyutin et al., 2021), (Zeng and Lu, 2022) or simple machine learning methods (Gil, 2022)). So far, no research has been found to analyse a one-minute time-term dataset on Gold using Neural Networks. According to (Weng et al., 2020) the annual growth rates of gold is 7.69% which is more than 6.79% of S&P 500 from 2014 to 2019, and their paper stated that gold data could be utilised for training models.

In the last decade, Neural Network (NN) technologies gained popularity in financial markets’ price prediction. They have been extensively employed in the financial sector in areas such as portfolio optimisation, High-Frequency Trading (HFT), risk management and equity portfolio management. One of the primary aims of researchers is to use Machine Learning (ML) models, which can produce results close to the price in the financial market within a relatively short certain period of the chosen time frame or determine the moment before the most probable signifi-
cant price movement. The current trend in using ML models is to create auto-trading bots to produce predictions close to the abstract capabilities of the human brain but with more accurate computational predictions.

Most of the research papers reviewed use Recurrent Neural Networks (RNNs) (Vargas et al., 2017), (dos Santos Pinheiro and Dras, 2017), which are specifically designed for next-hour to next-day price prediction based on sequence data.

The aim of this research is to predict the price movement of the next minute using neural networks on the one-minute Gold prices (XAU/USD) Forex dataset. The experiments were done using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM). The results obtained have been compared against the state-of-the-art algorithms that predict gold prices on more extended time frame data. This research demonstrates how neural network models can handle massive datasets and extract information from transaction signals to determine the optimal market entry points based on predicted significant movements in the next minute in order to minimise overall transaction costs, which is vital for intraday trading. A comparison of algorithm performance on error metrics with regard to the reduction in these error rates attained during prediction is also performed. The models are fine-tuned and evaluated with better training parameters for improving to reach a price prediction close to the actual values.

The remaining sections of this study are structured as follows: section 2 investigates the literature review and state-of-the-art studies on the topic of intraday financial trading; the proposed method is presented in section 3; section 4 outlines the results and comparison of models; section 5 discusses the validation based on the numerical analysis; finally the conclusions are drawn in section 6.

2 LITERATURE REVIEW

During the unstable, volatility time, investors are more interested in the gold market. Consequently, the liquidity of different financial instruments based on gold spot prices is increasing. For intraday trading, liquidity is crucially important, and it is possible to apply intraday trading methods used for currency and stock markets on the gold markets. As a result, the same methods used for predicting prices in the financial instruments in liquid markets, such as Forex exchange rates, liquid 'Blue chip' stocks (Big-tech, mining companies or global banks), and integral stock indices (Dow Jones, S&P 500, NASDAQ-100, FTSE 100) are employed for the gold markets.

Researchers have used a range of methodologies to examine the volatility of the gold price and its relationship to the variables assumed to impact it. (Manjula and Karthikeyan, 2019), (Chandrabai and Suresh, 2020), (Makala and Li, 2021), (Chen, 2022) used popular regression algorithms, such as support vector machine, ARIMA, linear and Lasso regressions, random forest, and gradient boosting, for gold datasets based on hourly and daily prediction. (Milke et al., 2020) used CNNs to predict the most probable significant price movements by analysing Forex tick data and measuring the performance metrics of the model.

(Manjula and Karthikeyan, 2019) validated different periods of data on a monthly basis to predict prices using three regression models: Linear Regression (LR), Random Forest (RF) and Gradient Boosting Regression (GBR) algorithms. To assess the quality of the prediction models implemented, they used the Mean squared error (MSE), Root mean squared error (RMSE) and Mean absolute error (MAE). Another research by (Chandrabai and Suresh, 2020) studied gold price prediction using Linear Regression, Random forest and Support Vector Machine (SVM) with R-Square and RMSE scores for evaluating the results. The same dataset on gold was used by (Makala and Li, 2021) in their research that studied ARIMA and SVM models on daily data from World Gold Council between 1979 and 2019 and measured the performance of the proposed models using RMSE and MAPE. The research results indicate that the SVM outperforms ARIMA in terms of MAPE and RMSE, with RMSE of 0.028, MAPE of 2.5 for SVM versus 36.18 and 2.897 for ARIMA. Due to SVM’s high accuracy, the findings imply that it should be employed for commodity price prediction. With the help of price components that possibly affect gold prices, (Chen, 2022) created an ensemble technique that combines SVM and LSTM models with quotient space theory.

The key purpose of the intraday trading is to gain instant profits from volatile prices of stocks, futures contracts and other financial instruments. Intraday trading can be defined simply as buying and selling of the same security during one day within one trading session (U.S. Securities and Exchange Commission, 2022). All positions should be sold before the end of the trading session, and no securities (including short positions) should be held overnight in order to avoid significant risks of getting losses from overnight gaps triggered by economic or political events (Shen, 2021).

Prediction of the next day’s price with some probability has become possible by searching for patterns...
of behaviour of market participants in historical data using complex neural networks and machine learning algorithms such as RNNs, CNNs and Gradient Boosting algorithms in regression models in conjunction with some Natural Language Processing (NLP) technologies. Some models use hybrid joint analysis approaches on input market data (prices and some technical indicators) together with news. The papers presented based on daily (dos Santos Pinheiro and Dras, 2017), (Vargas et al., 2017) and hourly price prediction (Madan et al., 2015) show that better accuracy is achieved on intraday data in conjunction with the news of the respective day, as previous days’ news had minimal impact and became noise. In their paper, (Vargas et al., 2017) employ Deep Learning (DL) algorithms to predict S&P 500 index intraday prices, using a headline technical indicators set and financial news as input. When using DL algorithms, it is possible to recognize and analyze complex non-trivial patterns and interactions of data, which are used to increase capital returns and automatically accelerate the trading process. (Vargas et al., 2017) focused on designs like CNN and RNN, and achieved significant results in typical natural language processing procedures for financial information. The results show that CNN performed better than the RNN for capturing semantics from text, and RNN is recommended to predict the stock market by modelling contextual information and complex temporal characteristics.

Using NLP techniques, (dos Santos Pinheiro and Dras, 2017) explored character-level language model pre-training with RNN for both interday and intraday forecasts of the S&P 500 index. This method outperforms other recent approaches in predicting the direction of the S&P 500 index for both individual shares and the index as a whole, demonstrating the high impact of current news on stock price movements in recent years.

Based on the hourly time-term dataset, the prediction method proposed by (Madan et al., 2015) uses minutes and seconds data of bitcoin to predict prices for the next 10-minute prices. Properties to predict the signs of future changes are modelled as bionomic classifications that are experimented with random forests and linear models. Their results were 50 to 35\% accurate in predicting signals of future price movements using a time frame of 10 minutes. Because there are many microscopic fluctuations and perturbations in the bitcoin price, as well as inside the prices of other financial instruments, 10-second intervals of data are used for a deeper understanding.

In their paper, (Zhang et al., 2021) applied hybrid LSTM models to predict the EUR/USD price movement. The classifier determines the direction of price movement as no action, increase or decrease in price. The model predictions included periods of one day, three days, and five days ahead. Results show that hybrid models outperform individual models for daily data.

In recent years, LSTM has been the most popular method employed for price prediction using time-series data. (Lim et al., 2019) used enhanced deep neural network LSTM to estimate the price value above or below in the next time step. (Siami-Namini and Namin, 2018) also used LSTM to forecast time series data for 12 stocks. The empirical findings in this paper show that DL-based algorithms, such as LSTM, outperform classic algorithms, such as ARIMA models. Similar results were published by (McNally, 2016).

Most research papers reviewed used daily close prices data as inputs for Neural Networks, Regression methods or a combination of networks. The authors identified a lack of research on intraday data, such as 1-minute and 5-minute time frames. Given the higher predictive performance of neural networks compared to classical machine learning methods and statistical time series analysis, this research focuses on modelling with neural networks, such as one-dimensional CNN and LSTM, which are also very popular for time-series data. The main contribution of this research consists of comparing and analysing predictions of two neural network models based on the one-minute gold Forex dataset for intraday trading to close the existing research gap.

3 PROPOSED METHOD

The expediency of short predictions of financial instrument’s price for the next second or minute can be explained by a simple logical statement that follows from the concept of causal determinism in economics (Chen et al., 2018) and is vividly described by the conception of Laplace’s demon (Johnson, 2017): if it is possible to predict the next second or minute, it is possible to predict the future any time ahead.

The prediction of the gold price for the next minute can show a potential acceleration of price change during the longer period inside the trading session. In other words, it shows the starting point of a possible momentum of future significant price change. Thus, from a practical point of view, a trading system based on this minute-by-minute neural network prediction only pays attention to the prediction of relatively big price changes in the next minute and does not react to a casual flat.

For this research the one-dimension CNNs has
been chosen because they are effective at time-series analysis (Chollet, 2016) and for pattern recognition in images. According to the literature review, LSTMs are widely used for time-series analysis as a benchmark. Therefore, it has been chosen as a second architecture for this paper.

3.1 Dataset

For the models to produce good results, it is crucial to choose a dataset that is error-free and detailed enough to reflect real short-term liquidity. Liquidity, in its most basic definition, is the quantity of cash and readily convertible assets a company has in order to meet its short-term debt obligations. In terms of trading, liquidity means the ability to sell a position in stocks, futures contracts, or other financial instruments without significantly affecting the current prices of these assets. According to the World Gold Council (WGC) (Das et al., 2022), gold is the second most liquid asset. Figure 1 demonstrates the changes of the gold prices from 1974 to 2022 (Dukascopy Swiss Banking Group, 2022).

Gold is one of the most liquid financial assets because it is traded with large volumes on the spot, futures, and Forex markets (Hundal et al., 2013). Therefore, evaluating and comparing various models using gold datasets is a promising resource for future research conducted on minute-to-minute datasets.

This research uses the gold market financial instrument (XAU/USD) in the Forex market. This one-minute interval dataset is publicly available on the Dukascopy Bank official website. Data from 2020 has been chosen because it has been extremely volatile due to the start of COVID-19 pandemic. The features include the trading volume and the Ask and Bid Values prices with date and time. There are 355,590 rows and 11 columns — Open, High, Low, Close, Volume values for both Ask and Bid, and local time for each minute.

The volumes of each transaction in the Forex market, including gold instruments, contain only some volumes bought and sold at current prices in the world since the Forex market is not centralised. However, the volumes of transactions indicated in this data repository sufficiently reflect the activity of market participants and increase when the market is activated. Thus, transaction volumes are an additional linearly independent parameter that has to be analysed (Milke et al., 2017).

3.2 Experiments

3.2.1 Convolutional Neural Networks

The essential step in the data preprocessing for time-series data is setting the rolling window size before training the neural network models. The 50-minute wide rolling window was used to generate mock images as inputs to the CNN. This window contains minute prices and volumes vs time, advancing in minutes left-to-right (increasing time) increments. Each minute generates a new chart for the previous 50 minutes, which then adds up to a 3D input tensor. Further, to reduce the size occupied in RAM and improve the metrics of correlation recognition, this 3D tensor was transformed from pictures into a 3D tensor of time series slices. The process described above is shown in Figure 2.

The 50-minute wide rolling window was chosen as the maximum possible based on the performance of the available GPU processors. For example, a 200-minute wide rolling window increases the number of input data by four times, and the number of calculations for training the neural network is proportional to 16 times.

As a consequence of this conversion, 2D arrays of prices and volumes are converted into a 3D tensor. The MinMaxScaler technique is used to normalize the data that is then fed into a CNN for training and pattern recognition. The sliced 2D windows in the 3D tensor of inputs.

The 50-minute wide rolling window was chosen as the maximum possible based on the performance of the available GPU processors. For example, a 200-minute wide rolling window increases the number of input data by four times, and the number of calculations for training the neural network is proportional to 16 times.

As a consequence of this conversion, 2D arrays of prices and volumes are converted into a 3D tensor. The MinMaxScaler technique is used to normalize the data that is then fed into a CNN for training and pattern recognition. The sliced 2D windows in the 3D tensor of inputs.
The tensor may be associated with the previously known financial practice of analysing Japanese candlestick patterns, but on a newer technological platform (Nison, 2011).

After preprocessing, the data is divided into training (including 20% for validation) and testing datasets in a ratio of about 1:10; approximately 230 days for training and 30 days for testing.

The 2-convolutional layers CNNs, with 50 and 100 neurons each, are trained. Then max-pooling layer is added to shrink the dimension with the aim of reducing the redundant characteristics. After the 2-layer CNN network, a dense layer with 25 neurons, with ReLU activation, was added to allow for additional customisation. Then the sequence proceeded to the fully connected layer with the softmax activation function to forecast expected outcomes; a dense layer of four neurons was chosen as the last layer for categorising samples into Open, High, Low, and Close values. Overall, 40,500 total weight parameters were optimised during training.

The above conventional neural network architecture was chosen to receive preliminary training outcomes to assess the approach’s perspective. In further research, the authors intend to use AutoML tools to determine the optimal wide of the rolling window and the best hyperparameters of the neural network used.

### 3.2.2 Long Short-Term Memory

The data preprocessing for the LSTM consists in reshaping the data to a one-dimensional array from the CNN model’s close price data. This analysis can be considered a regression task. Similarly, the train validation and test split of data are kept the same.

The first experiment used a single-layered LSTM network where the 64-neuron layer is followed by a 25-neuron dense layer and a 4-neuron output dense layer.

In the second experiment, a two-layered LSTM network was built. It has 64 neurons on the first LSTM-layer, followed by 128 neurons in the second LSTM-layer, and the rest is the baseline of the dense layers described above. Though the computational time for two-layered networks is slightly higher than for one-layered networks, the loss is 100 times less.

Finally, a third LSTM-layer with 256 neurons was added to the two-layered LSTM network, which has performed better both in specific computational speed as well as with test loss. During training, a total of 516,400 total weight parameters were tuned.

### 4 RESULTS AND COMPARISON

Experiments with CNNs are carried out with different convolutional layers, optimisation algorithms, loss functions, and metrics. 2-layer ReLu CNN networks were compared with 4- and 6-layer networks.

Keras Model Checkpoint callback function was utilised to remember training results for every epoch, with the epoch with the best error metrics being chosen as the subsequent choice. Tests were conducted with data batch sizes ranging from 10 to 100 and epoch counts ranging from 5 to 100 for this research. During the analysed period, 20 data batches and 10 to 50 epochs which were determined by the above callback function, produced the best results.

The CNN with the parameters showed in table 1 demonstrates a good performance.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss Function</td>
<td>MSE</td>
</tr>
<tr>
<td>Optimizer</td>
<td>Adam</td>
</tr>
<tr>
<td>Batch Size</td>
<td>20</td>
</tr>
<tr>
<td>Number of neurons in each layer</td>
<td>50</td>
</tr>
<tr>
<td>Epochs</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 2 presents the results of both CNN and LSTM for the error metrics MSE, MAE and MAPE, used to evaluate model efficiency on each epoch. The decrease in the error values shows that the model performs well on the training and validation dataset.

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>MAE</th>
<th>MAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>0.000001017</td>
<td>0.001208</td>
<td>11.55</td>
</tr>
<tr>
<td>LSTM</td>
<td>0.000013088</td>
<td>0.002531</td>
<td>60.33</td>
</tr>
</tbody>
</table>

The predicted prices produced by the CNN on the test data from 2020 are shown in Figure 3, alongside the actual values of Open, High, Low and Close.

The visualisation for the LSTM training process is shown in Figure 4 with Training and validation losses. Figure 5 demonstrates the LSTM model results of Train, Test, Valid and Prediction prices.

Following the outcome of the models developed, the CNN was chosen for further experiments as its performance is greater than the one of LSTM for the dataset used. After changing and comparing the NN architectures and hyper-parameters, the CNN was trained with parameters that are performing better to produce fewer error rates.
5 DISCUSSION AND CRITICAL ANALYSIS

The ability to recognise financial market patterns has clear benefits for investors. For this paper we employed the CNN and three-layered LSTM models to forecast the future movement of intraday financial market transactions utilizing non-standard approaches for preprocessing raw data. The normalisation of nonlinear data as well as the conversion of 2D data into a 3D tensor via the production of consecutive data are both essential components of the approach’s design. As a consequence, the size of the incoming data grows substantially and impacts the available resources. In this research, the data was normalised using the min-max approach that is one of the most often used data normalisation strategies.

In transactions involving very high volume lots, significant price fluctuations and probable delays in the execution of market orders often occur, decreasing the prediction’s quality. Therefore, deep learning techniques, such as CNN and LSTM were chosen as they are more accurate than basic neural networks.

Preliminary findings indicated a reasonably high accuracy and error rate, which should yet be thoroughly confirmed in calculations, taking into consideration the constraints imposed by the time-series data set utilized in the calculations. The data was separated into training and validation datasets in a 9:1 ratio, with 90% of the dataset utilized for training and 10% for results validation. The results found that baseline 2-layered CNN with the parameters shown in table 1 had the best performance in terms of MSE, MAE, MAPE. The decrease in the error values shows that the model has performed well on the training and validation dataset.

After changing and comparing the hyper-parameters, the model was trained again and it produced fewer error rates.

6 CONCLUSION

In this research, two neural network architectures were applied to gold Forex market data based on gold spot market prices, using various deep learning approaches and non-linear data preprocessing. These approaches were utilised to predict the intensity and volatility of price movements in a short-term period.

This research presents a CNN model for predicting future intraday financial market price accelerations utilising non-standard data processing approaches. Based on the test results, it can be concluded that the primary contribution of this research is a statistical confirmation of the high probability of the ability of traders (natural or legal persons) to make profits through intraday forecasting of gold
price movements.

On the basis of various hyper parameter values, we compared the performance of the CNN and LSTM neural networks in forecasting the prices of the gold market index in intraday time periods. To evaluate the resilience and performance of these models, we compared the error metrics of the two neural networks and concluded that CNN performs better based on low error rates.

Future research should focus on including AutoML to optimise Hyper-parameters with hyper-parameter optimization (HPO), Reinforcement Learning by normalising raw data. HPO eliminates the need for a human expert to perform the time-consuming task of hyper-parameter tuning. It is possible to create self-learning agent using reinforcement learning (RL) with the recognises not only market entry points with the highest probability of profit but also market exit points. Such a system will be evaluated in terms of maximising the Sharpe ratio with RL algorithms like Proximal Policy Optimization (PPO), Q-learning and, Deep Q Neural Network.

REFERENCES


Raimundo, M. S. and Okamoto Jr, J. (2018). Application of hurst exponent (h) and the r/s analysis in the class-