# A Spatiotemporal Forex Trading System Based on a Hybrid Model GAT-LSTM: Forecasting Forex Price Directions

Nabil MABROUK, Marouane CHIHAB, Younes CHIHAB
Computer Sciences Laboratory, Ibn Toufail University, Kenitra, Morocco

***Abstract*—Due to the high volatility and complex interdependencies within financial markets, predicting Forex prices becomes a difficult challenge for investors. Furthermore, the traditional trading models struggle to capture those relationships. To address this issue, we introduced a spatiotemporal Forex trading system, GAT-LSTM-based; it is a hybrid approach that combines Graph Attention Network (GAT) with a Long Short-Term Memory (LSTM) network. The GAT component helps to capture spatial dependencies between currencies by constructing a directed graph containing 28 currency pairs alongside commodity stock and US stocks. The strength of the GAT component lies in its ability to dynamically adjust and recalculate the weights of edges over time, which helps our proposed system to adapt to macroeconomic changes, news events, and financial factors that can impact the Forex market status. The LSTM component deals with the nature of time series datasets. It learns temporal interderdependencies, allowing our system to detect repeated long-term patterns over time. Experimental results proved that the suggested hybrid model, GAT-LSTM, surpasses both LSTM and GAT separately. By combining both elements and leveraging simultaneously the strength of dynamically modelling spatial dependencies, and the strength of learning long-term temporal patterns, our suggested system became more accurate in forecasting Forex price directions, showing promising results and high accuracy during the validation phase.**

***Keywords*—*Forex trading; hybrid deep learning model; Graph Attention Network (GAT); Long Short-Term Memory (LSTM); spatiotemporal forecasting***

## I. Introduction

The Forex, which is a short abbreviation of foreign exchange market, is known for its high volatility and strong price fluctuation caused by the influence of economic news, investor sentiment, and global events [1], [2]. Thanks to Forex advantages, including high liquidity, low transaction costs, 24-hour availability/operations, and online trading possibilities [3], investors (companies, individuals, and institutions) consider Forex as a profitable market that offers a good opportunity to achieve their goals and specific financial objectives utilizing exchange rate volatility [4], [5]. Most currencies are traded against the USD, and they are traded in pairs [6]. Each pair is composed of base and quoted currency [7]. Traders can exchange currencies at any time because the Forex market operates continuously for 24 hours a day, five days a week, which leads to elevated liquidity and a substantial trading volume [8], [9]. The historical data of Forex transactions has become the fuel of research.

In the past, numerous studies have been conducted on predicting Forex prices. Investors have deployed and adapted traditional trading strategies and other tools that they found helpful in their previous operations to suit their trading styles [10]. Three main techniques of forecasting Forex prices for time series are adopted by traders and investors: The first is fundamental analysis, which involves examining economic factors, political characteristics, and news events that may influence the Forex market status [11], [12]. The second is technical analysis (TA), which is a trading method based on historical data and mathematical calculations. To get an idea about future price variations, technical analysts look for past behaviors and patterns in historical data, using technical indicators and charts. The third type of analysis is sentiment analysis. Forex market fluctuations are primarily influenced by emotions and sentiments [13]. It studies the general mood of the market and analyzes the general attitude of traders toward the Forex market. Forex prices fluctuate because of investors' emotions, whether they are optimistic or pessimistic about the market. It is the primary factor for their decision-making to buy or sell. Trading techniques follow technological advancements.

Traders use computational algorithms, specifically deep learning, data mining, and machine learning, to create their trading strategies [14]. However, simple machine learning and deep learning algorithms struggle to forecast Forex prices accurately because of the social, economic, and political interdependencies.

In this study, we propose a hybrid model to leverage the dependencies between Forex currency pairs, the commodity stocks, and the US-stocks. Our spatiotemporal system can automatically forecast and identify future price directions, trends, and cycles, leveraging the time series nature of the Forex dataset. The spatial part of our model will deal with horizontal dependencies between currency pairs and stocks, forming a Graph to overcome the issue of economic, social, and political factors; each node of this Graph is enriched with valuable features such as technical indicators and trading rules to improve the prediction accuracy. Besides forming a Graphical representation, and thanks to the GAT component, which can update the strength of those connections between nodes, our model is able to adapt to global changes. The temporal part is to deal with the vertical dependencies between Forex transactions by learning and capturing long-term patterns. Our solution can perform the optimal trade decision, including buy, hold, or sell signals.

## II. Related Works

In the last decade, forecasting currency exchange rates has been the center of much research; many techniques and methodologies have been adopted in this field of research. Technological development helped investors trade at any time and from any place around the world through electronic platforms. [11], leading to the accumulation of a large set of historical data; which motivated researchers to conduct studies on the resolution of financial market problems based on computational algorithms [15]. In [16], Neely and Weller's research examines the manner in which technical analysis is employed in the foreign currency market, highlighting its widespread acceptability and superiority compared to fundamental analysis, particularly in short-term trading scenarios. Initially observed by traders in the London market, their findings have been validated across various market conditions, underscoring the enduring significance of technical analysis in Forex trading. In [17], they provide significant insights into the fundamental concepts of technical analysis, emphasizing its reliance on historical stock price and volume data to discern market patterns. Unlike fundamental analysis, which evaluates a stock's intrinsic worth, technical analysis emphasizes chart patterns and statistical indicators to forecast future market trends, offering a distinct perspective on the evaluation of equity assets. In [18], they conducted research on using SVM for market trend prediction; the results showed an out-performance of SVM with Expert Advisor, versus a simple Expert Advisor. They used an SVM (support vector machine) to forecast trends, "UP trend" or "DOWN trend". In [19], they showed that KNN performed better using PCA. KNN can reduce redundant and noisy information, leading to accurate predictions of the stock and Forex markets.

In [20], Zhelev and Avresky assert that the referenced literature on deep learning helps as a fundamental basis for handling the complex problem of Forex price prediction. In recent years, deep learning has shown notable progress in resolving Forex trading tasks and other domains [21]. Deep Neural Network algorithms have been employed to forecast time series data and quantitative trading [22], [23], [24]. Deep neural network algorithms have shown improved performance and accuracy, producing better returns in the finance sector [25], [26]. According to Galeshchuk [27], an ANN model with a single hidden layer can outperform the time series model in forecasting the price, especially for short-term prediction of the market. Hu et al. [28] applied NLP techniques in news to select only authentic content and influence trend. They proposed a hybrid LSTM model with an Attention mechanism for the prediction task. In [29], they confirmed that ANN performs better than DNN and SVM in time series data. In [30], the authors indicated that, while methodologies utilizing neural networks demonstrate significant potential, they are frequently evaluated solely on the basis of their prediction errors. However, it is equally imperative to assess their impact on a real-world trading system, where the decision to invest in the market is as critical as the decision not to invest.

In [31], they proposed time series data mining based on GNN and transformer to tackle the problem of modeling complex dependencies, using an adaptive adjacency matrix. They showed that the model can dynamically calculate the relationship between variables and capture local dependencies. For a multivariate time series forecast, in [32], they developed an architecture that merges a spatial-temporal graph neural network with a filtering module. They forecast future sales from an artificial time series sales volume dataset. The suggested spatial-temporal model showed excellent performance compared to baseline approaches without graphical architecture. In [33], they used a graph-awareness-based GAT-GAN Generative Adversarial Network to learn temporal dependencies that capture spatial relationships and generate long-time series data. In [34], they combined long-short-term memory (LSTM) with graph attention networks (GAT), creating a spatiotemporal model to accurately predict the water quality of the Pearl River Basin. They confirmed that their model showed robust performance and succeeded in jointly capturing spatial and temporal interdependencies. In [35], the authors examined the influence of human movement on the transmission of COVID-19. It establishes a mobility network that delineates travel patterns among Brazilian cities and incorporates time series data. Two Graph Convolutional Network (GCN) neural networks are employed to forecast time series for each city, considering adjacent cities. The study assesses LSTM and Prophet models, concluding that GCN models produce more consistent results than the assessed models. The Prophet model exhibits a greater maximum RMSE value.

## III. Our Proposed Solution

### A. Methodology and Problem Definition

The global economy is interconnected, a phenomenon known as internationalization, causing the interdependency of multiple financial sectors. However, Forex refers to the exchange of currencies. This fact implies the existence of dependencies between currencies. We assume that each pair of currencies is related to other pairs. Forming a directed graph G = (V, E) [36], where the set of nodes V represents all currency pairs, linked by an edge set E. these connections and the structure of the graph G are represented by a matrix adjacency A of two dimensions (D x D), where D is the number of nodes ( in our case it is currency pair ); each element $A_{ij}$ represents the existence of an edge between two nodes $V_i$ and $V_j$ [see Eq. (1)]:

$$A_{ij} = \begin{cases} 1, & \text{if an edge exists between } V_i \text{ and } V_j \\ 0, & \text{Otherwise} \end{cases} \tag{1}$$

The Graph Neural Network (GNN) uses an adjacency matrix to capture spatial dependencies between currency pairs and to learn deep complex patterns. Providing the necessary information for accurate predictions. The GAT-LSTM model is trained using a supervised learning approach, which consists of minimizing the disparity between the forecasted and actual trends in historical data. The backpropagation during the training process allows the model to learn the optimal weights for both the GAT and LSTM components.

### B. Data Description

*1) Data collection:* This study aims to predict the variation of prices in the Forex market. For that, we obtained the necessary dataset from various reliable financial platforms such as ForexSB, FXStreet, DailyForex, etc., which offer valuable historical Forex data for time series analysis.

To construct comprehensive data for our model, we considered the 28 most common currency pairs. They can be classified into two general classes: the first is major currencies (GBPUSD, EURUSD, USDCHF, USDCAD, and USDJPY), and the second class is the minor currencies (AUDCAD, NZDJPY, NZDUSD, CADCHF, NZDCAD, GBPNZD, NZDCHF, GBPCHF, GBPJPY, GBPAUD, AUDNZD, EURNZD, EURJPY, AUDCHF, EURGBP, EURCHF, EURCAD, EURAUD, CHFJPY, CADJPY, AUDUSD, GBPCAD, AUDJPY).

The objective is to forecast future price directions for the five major currency pairs. In addition to the 28 currency pairs, we used six additional stock data, such as commodity stocks and US stocks, to enhance the predictive performance of our algorithm. The collected dataset covers the period from January 2009 until June 2025, spanning 5006 days, resulting in approximately 170,204 daily observations (5006 days x 34 stocks). This dataset provides a detailed view of the exchange rates for the currencies, presented by five features:(Open, High, Low, Close, Volume) for each stock. Where Open is the price at the beginning, High is the highest price, Low is the lowest price, Close is the final price, and Volume is the total amount of currency exchanged. Fig. 1 illustrates the scaled closing prices of the five major currency pairs.

*2) Data preprocessing and feature engineering:* In the data pre-processing phase, we handled the missing values to avoid errors in prediction models by ensuring that there were no missing values in our time series dataset. In addition to primary features, we generated two other variables, ”return” and ”direction” features, that indicate whether the prices will increase or decrease in the following sequence [see Eq. (2)]:

$$Direction = \begin{cases} 1, & \text{if Price }_{t+1} > \text{ Price }_t \\ 0, & \text{Otherwise} \end{cases} \tag{2}$$

To improve the performance of our model, we created additional features that are generated using mathematical calculations. Those features are called technical indicators, such as:

- The Bollinger Band (BB),
- Moving Polynomial Trending (MPT)
- The Relative Strength Index (RSI),
- The Commodity Channel Index (CCI),
- The Weighted, Exponential, Simple, and Convergence Divergence Moving Average (WMA, EMA, SMA, and MACD),
- The average directional index (ADX),
- The Rate-of-Change (ROC)

Technical indicators can be incorporated as additional features, enriching our dataset with additional information that helps to improve the accuracy of the prediction. We finally form a comprehensive spatiotemporal 3D matrix for model training and evaluation (5006 days x 34 stocks x 81 features).

*3) Feature matrix X:* Let X be a feature matrix of dimension (N x P), where N represents the total number of currency pairs and P is the total number of features of each currency pair. The features matrix $X \in \mathbb{R}^{\mathrm{N \times P}}$ is defined by the following Eq. (3):

$$X^k = \begin{bmatrix} x_{11}^k & \cdots & x_{1p}^k \\ \vdots & \ddots & \vdots \\ x_{n1}^k & \cdots & x_{np}^k \end{bmatrix} \tag{3}$$

Our dataset is a 3D matrix, where $X^k$ represents a feature matrix at row k.

*4) Feature scaling:* During the training process, features with larger magnitudes dominate, preventing the model from learning. To optimize the efficacy of our model and guarantee that all features contribute equally to the training process, we have scaled the data using the Min-Max scaler in the range [0, 1]. See Eq. (4):

$$X_{sc} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{4}$$

where,

Xsc: the scaled value of the data point.

X: the initial data point’s value.

Xmin: the minimal value of the data point.

Xmax: the maximum value of the data point.

*5) The spatial representation of currency pairs (Graph):* We assume that the currency pairs form a graph. Each currency pair is considered a node; the relationships between those nodes (edges) are determined based on the calculation of correlation coefficients, forming a weighted graph. Those weights are dynamic and change during training, enabling our model to self-adjust with global factors such as economic cycles, news events, and industry trends. Initially, we created a fully connected graph: Every currency pair is connected to every other currency pair. Fig. 2 illustrates the graph representation of currency pairs.

*6) Sequence creation:* Our objective is to improve the predictive capabilities of GAT and LSTM by effectively leveraging their combined power. We reconstructed the resulting data into sequences by implementing a sliding-window approach.

Let $X = \{x_1, x_2, ..., x_T\}$ be a time series dataset, where $x_t$ represents the data point at time t, and T is the length of the initial dataset. The $i^{\text{th}}$ sequence $S_i$ of length L is defined in Eq. (5):

$$S_i = \{x_i, x_{i+1}, \ldots, x_{i+L-1}\} \tag{5}$$

For example, a window of five days length means that to forecast the price direction Yt of a currency pair, we feed our model with the data that correspond to time step t, along with the data of the four prior time steps [$x_{t-3},\ x_{t-3},\ x_{t-2},\ x_{t-1},\ x_t$]. This approach helps the LSTM component to better detect temporal dependencies. Furthermore, for better optimization of the LSTM layer, we have trained our model using various sequence lengths between 1 and 25.

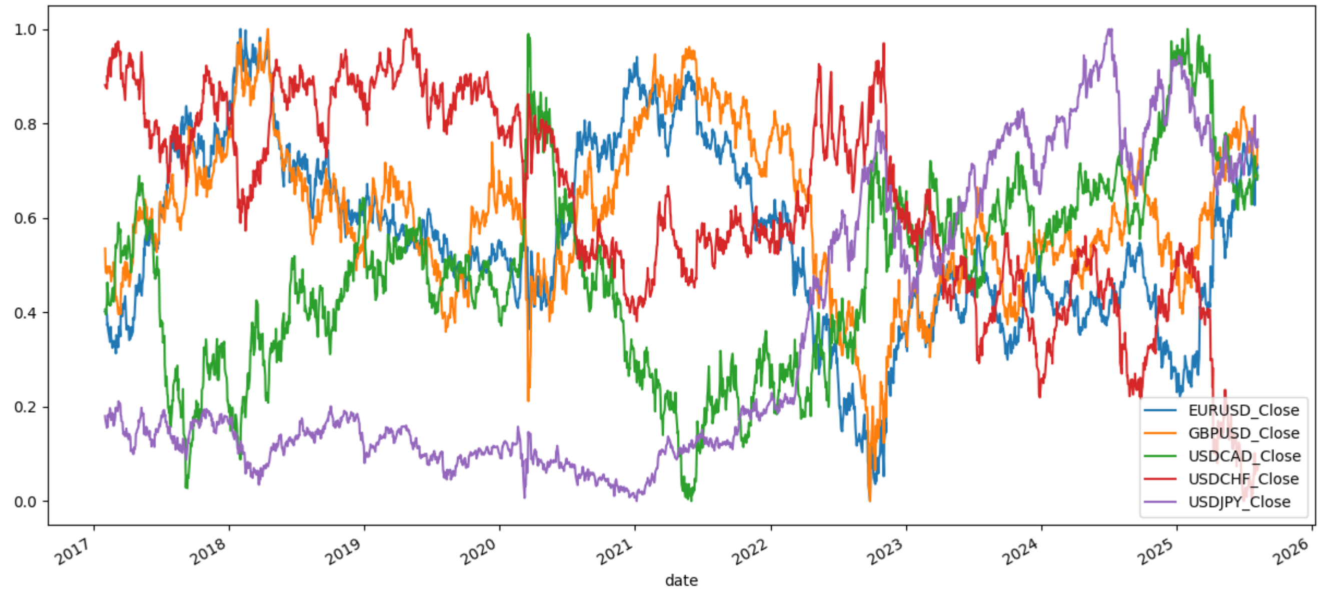


Fig. 1. The scaled closing price of the five major currency pairs.

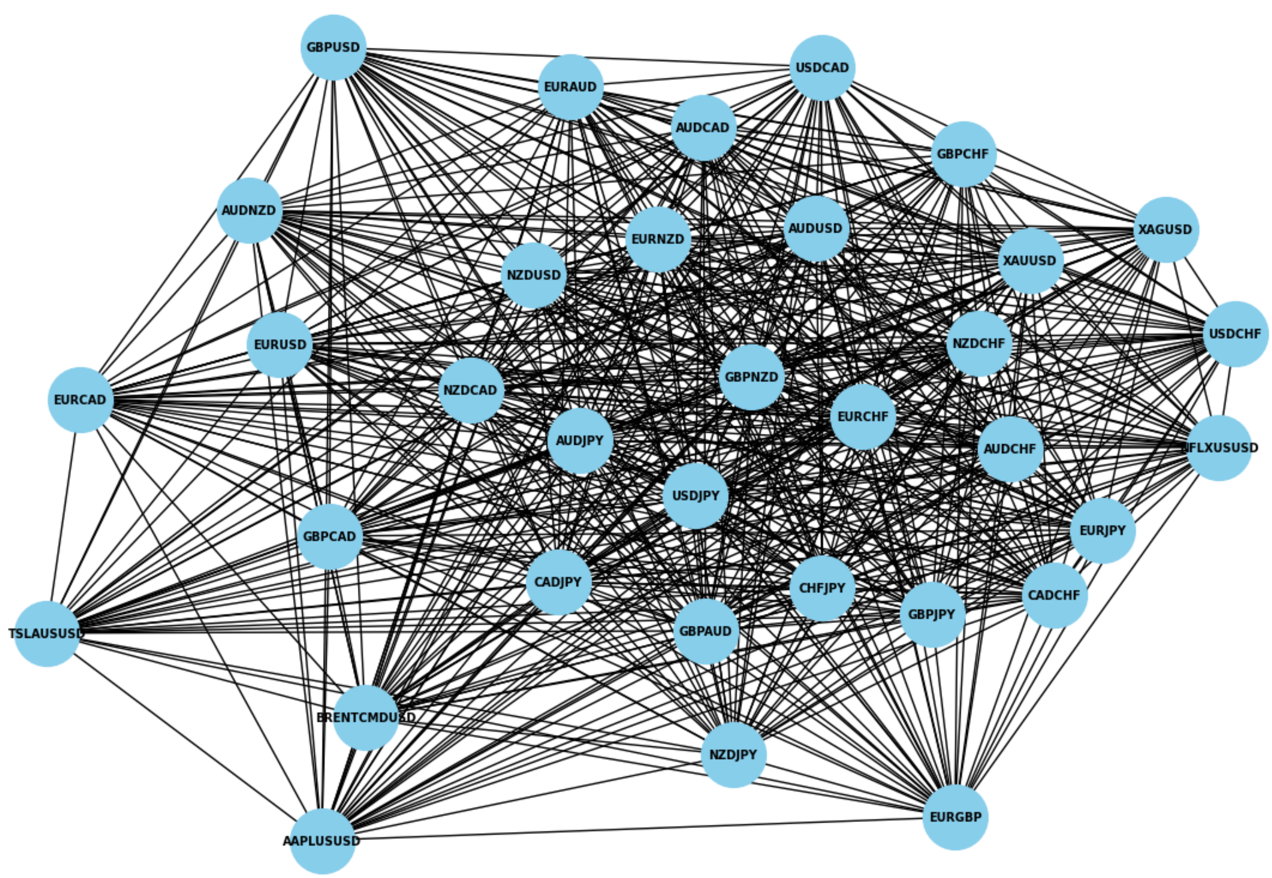


Fig. 2. A complete graph with 28 currency pairs and six stocks.

*7) Split data into train and validation:* To ensure a realistic and rigorous evaluation, we implemented a strict chronological data split to mimic real-world prediction. The dataset was divided temporally as follows:

- Training set (70%): January 1, 2009, to December 31, 2019.
- Validation set (10%): January 1, 2020, to December 31, 2021.
- Test set (20%): January 1, 2022, to May 30, 2025.

Crucially, we employed a strict procedure to prevent data leakage. All preprocessing steps were calculated solely from the training data, and those transformations were then applied to the subsequent validation and test sets. By maintaining this chronological order and separation, we were able to accurately assess the capacity of the model to generalize to future unseen data.

## C. Model Architecture

In this study, we implemented two independent models, the first is Graph Attention (GAT) Network and the second is Long Short-Term Memory (LSTM) Network, to forecast Forex trend direction. We compared those models with a hybrid model, GAT-LSTM, which is a spatiotemporal model that combines two components to take advantage of the strengths of both architectures in dealing with spatial and temporal interdependencies simultaneously.

*1) GAT (Graph Attention Networks):* Like any Graph Neural Network (GNN) architecture, the Graph Attention Network (GAT) is designed to analyze and understand graph-structured data; in addition, it implements the attention mechanisms [37]. GAT assigns attention weights to neighboring nodes. Updates the node representation according to the importance of the neighbor's information and refines the relationships between individual pairs of currencies. This aids in generating a dynamic, weighted graph, which subsequently helps manage fluctuations in the Forex market. The following Eq. (6) mathematically defines the attention mechanism:

$$\alpha_{\mathrm{ij}} = \mathrm{softmax}_{\mathrm{j}}\left(a\left(W * h_{\mathrm{i}}, W * h_{\mathrm{j}}\right)\right) \quad (6)$$

Let $\alpha_{ij}$ be the attention coefficient between currency pair i and currency pair j, where $h_i$ and $h_j$ denote the characteristic vectors of currency pair i and currency pair j, respectively; W is a learnable weight matrix, while a is an attention function. The attention mechanism dynamically adjusts the importance of each edge based on the learned attention weights. By leveraging attention mechanisms, GAT can effectively overcome the issue of stock variations caused by various factors.

*2) The Recurrent Neural Networks and the Gated Recurrent Unit:* To address sequential data challenges, including time series and text, researchers have created the Recurrent Neural Network (RNN) [38], [39], [40], a specialized neural network that incorporates a memory mechanism to retain information from prior inputs. Feedback links in their design enable the retention of knowledge across temporal stages. Recurrent Neural Networks (RNNs) excel in sequence modeling problems, where the temporal context of data points is essential for precise prediction and comprehension.

*3) LSTM (Long Short-Term Memory):* Long Short-Term Memory (LSTM) was published by [41], [42], [43], [44]. It is an architecture of Recurrent Neural Networks (RNNs) that can help address the problem of vanishing gradients faced by standard RNNs. LSTM can effectively bridge time lags exceeding 1,000 discrete time steps. This capability makes them suitable for tasks where long-range temporal dependencies are critical.

The idea behind using an LSTM layer in our hybrid model is to learn and recognize sequential patterns that characterize time series datasets and capture temporal dependencies of Forex prices. We have adopted a many-to-many architecture to forecast the following trend of the five major currency pairs.

*4) Hybrid model GAT-LSTM:* In this study, we have proposed a hybrid model to effectively forecast Forex trends by capturing spatial dependencies between currency pairs and learning sequential patterns simultaneously. This model, GAT-LSTM, is a combination of two components, GAT and LSTM; see Fig. 3. By integrating the GAT and LSTM networks, we succeeded in creating a spatiotemporal model that takes advantage of the strengths of the first component, GAT, in dealing with spatial dependencies in the Forex by defining edges between currencies. GAT helps adjust the weights of edges while learning and training, which enables our model to adapt to economic changes. The second component, LSTM, is designed to handle the nature of time series datasets and detect temporal patterns that are repeated over time.

The model procedure begins with integrating multidimensional features into a predefined directed graph. The GAT module applies attention mechanisms, allowing each node to selectively absorb more information from neighboring nodes based on relevancy weights. This capability enables the model to handle diverse inputs from various sources, which results in spatially enhanced feature representations. Then, these spatial representations are reorganized into chronological patterns to be used by the LSTM module for capturing dynamic patterns in Forex prices, such as trends and cycles. The strength of LSTM lies in effectively modeling both short-term variability and long-term dependencies, which are critical to accurately forecasting dynamic currency prices.

The final step is the fully connected layer (the output layer). It is a linear layer that takes the temporal features extracted by LSTM and generates predictions for the trend direction. The role of this layer is to perform a classification task; the output is a prediction of the direction of the currency price (buy, hold, or sell) of the five major currency pairs; the number of units corresponds to the number of predicted classes.

## D. Training Phase (Hyperparameter Optimization and Fine-Tuning)

For efficient training, we ran our model for 50 epochs and finetuned it by optimizing the hyperparameters during the learning phase; the batch size is around 24 and 32. And for loss function optimization, we have used the Adam optimizer with a learning rate around 0.001. Our algorithm attempts to resolve a classification issue by forecasting the Forex trend direction. We employed BCEWithLogitsLoss for the validation and training loss function, which is used for multilabel classification problems. BCEWithLogitsLoss combines the sigmoid

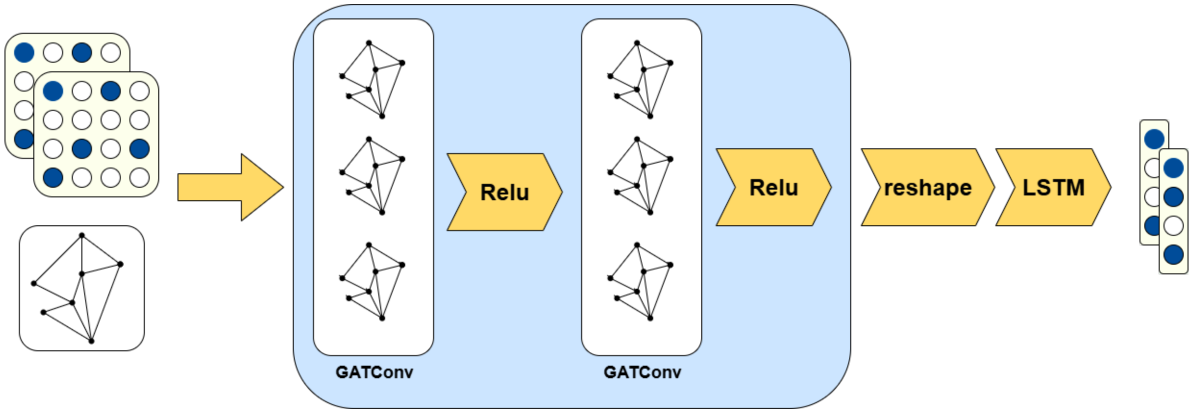


Fig. 3. The flowchart of the GAT-LSTM hybrid model for Forex price direction forecasting.

TABLE I. COMPARISON OF MODELS' PERFORMANCE (ACCURACY)

| | LSTM | GAT | GAT-LSTM |
|---|---|---|---|
| **EURUSD** | 0.6727 | 0.5206 | 0.8369 |
| **GBPUSD** | 0.6856 | 0.5026 | 0.8216 |
| **USDCAD** | 0.6134 | 0.4845 | 0.8267 |
| **USDCHF** | 0.6314 | 0.5077 | 0.8236 |
| **USDJPY** | 0.6418 | 0.5361 | 0.8287 |

and BCE loss into a single function; the formal definition is given in Eq. (7) as follows:

$$\text{Loss} = \text{mean}\left(BCE\left(\sigma\left(x_i\right), y_i\right)\right) \quad (7)$$

where, $x_i$ and $y_i$ represent the logits and true labels for the $i^{th}$ element in the batch, respectively. The mean() function calculates the average loss over all samples in the batch. The loss can also be configured to return the sum of the losses instead of the mean, depending on specific needs.

## IV. Discussion and Results

### A. LSTM Performance

The LSTM model is trained for 30 epochs using both architectures, with and without the window approach. In both methods, the model performed moderately in capturing temporal dependencies in currency price movements. The accuracy was between 61% and 67%, as shown in Table I, indicating that the LSTM model struggles to predict the general trend with specific turning points.

### B. GAT Performance

We trained the GAT model for 50 epochs; its performance is slightly lower than that of LSTM. The accuracy did not surpass 50% which is not impressive, see Table I. These results could be caused by the graph structure, which considers only horizontal dependencies between nodes and ignores the vertical dependencies characterized by the sequential nature of stock price data. The weak correlation between nodes was not enough to provide significant advantages.

### C. GAT-LSTM Performance

Compared to GAT and LSTM models, the hybrid model GAT-LSTM showed better performance in forecasting Forex trends as shown in Table I. We trained GAT-LSTM for 30 epochs, using different window size as described in Table II. After analyzing the training and validation loss graphics, as illustrated in Fig. 4 and Fig. 5, we found that the model is well-trained and fits the training data well, and it can be generalized for a larger time frame. We observed the lowest accuracy value (72%) at the lowest window size (1 day); while the highest accuracy value ( 83.7%) is observed at 20 days window length; the accuracy corresponds to the sequence length; When the window length is higher than 20 days, the algorithm becomes slower and less performant, the accuracy starts decreasing. We conclude that the optimal size of the sequence length is 20 days. The fact that the model's performance corresponds to the size of the window is because feeding the model with additional significant features helps identify patterns.

The high performance of LSTM-GAT is relevant to the fact that it is a hybrid model combining temporal modeling (LSTM) with relation modeling (GAT), which leads to better learning of Forex market dynamics. The LSTM is responsible for capturing individual Forex price patterns, while GAT captures the dependencies between currencies. That helps improve the predictions.

### D. Generalization and Performance During Conflict

While our primary focus is on demonstrating the effectiveness of the GAT-LSTM model for Forex price direction forecasting, we recognized the need to account for real-world trading frictions like transaction costs (spreads) and slippage. To provide a more realistic perspective, we applied a moderate 10% reduction to backtested profits achieved during the period test, specifically the COVID-19 pandemic. Even with this adjustment, our system exhibits strong potential for profitability and resilience during crises.

*1) Return on Investment (ROI):* We evaluated the performance of our model using the same loss function, BCEWithLogitsLoss, and calculating the accuracy of prediction as shown in Fig. 6 and Fig. 7. Furthermore, we back-tested our proposed system using financial metrics that consist of calcu-

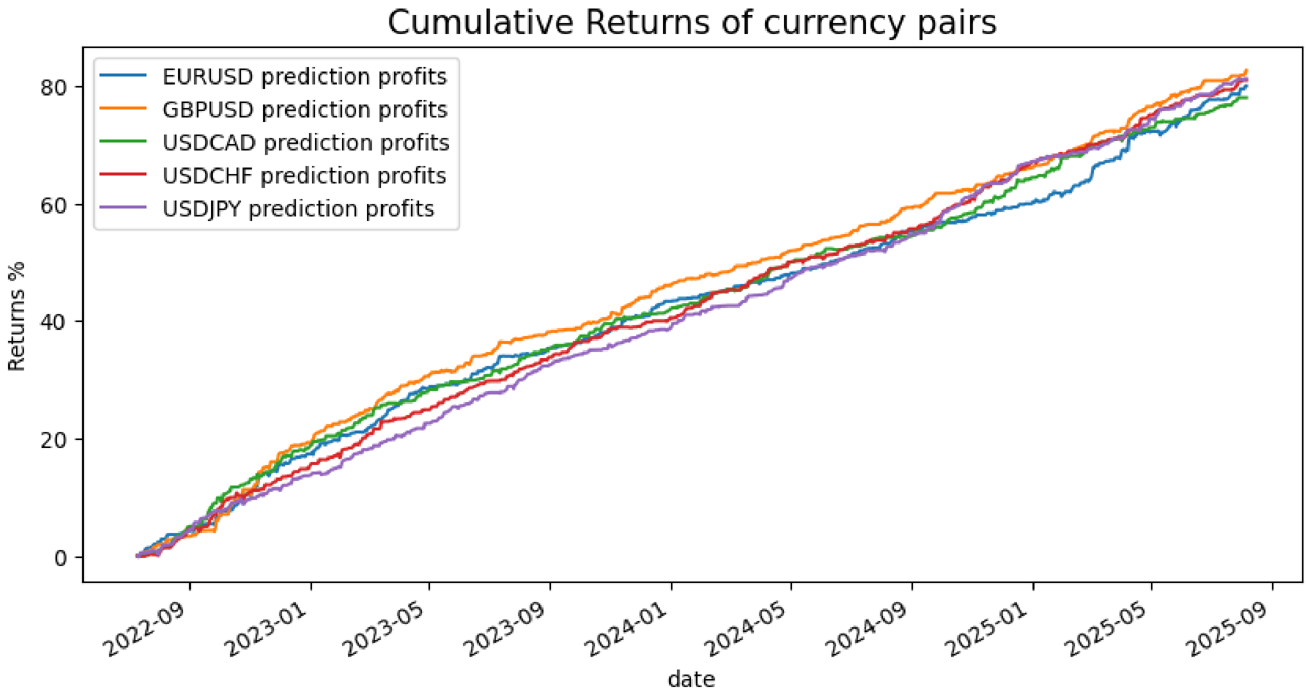


Fig. 4. GAT-LSTM training/validation loss using 20 days window size.

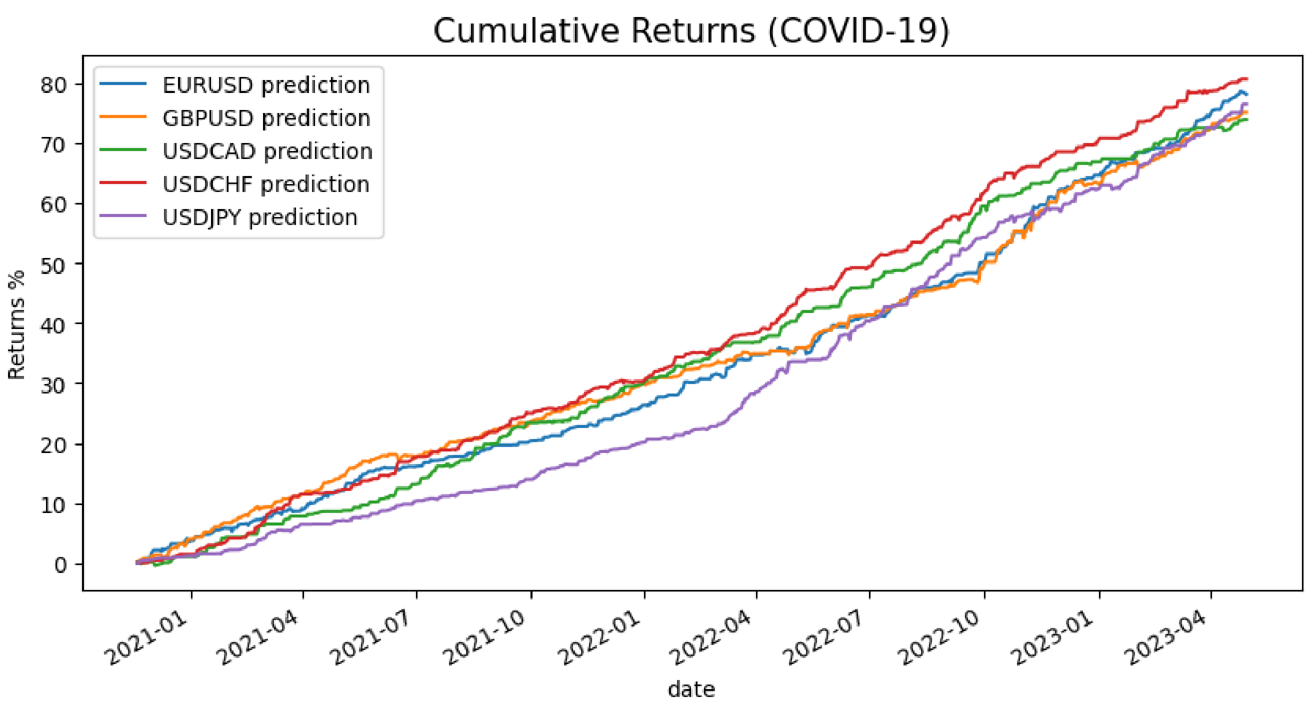


Fig. 5. GAT-LSTM training/validation accuracy using 20 days window size.

TABLE II. MODEL ACCURACY BY WINDOW-SIZE AND TIME FRAME

| Currency pair | Window size | | | | | |
|---|---|---|---|---|---|---|
| | 1day | 5 days | 10days | 15days | 20days | 25days |
| **EURUSD** | 0.7462 | 0.7686 | 0.8287 | 0.8165 | 0.8369 | 0.8217 |
| **GBPUSD** | 0.7339 | 0.7472 | 0.7910 | 0.7961 | 0.8216 | 0.8147 |
| **USDCAD** | 0.7309 | 0.7503 | 0.7910 | 0.8135 | 0.8267 | 0.8094 |
| **USDCHF** | 0.7523 | 0.7564 | 0.8114 | 0.8114 | 0.8236 | 0.8159 |
| **USDJPY** | 0.7288 | 0.7666 | 0.8063 | 0.8022 | 0.8287 | 0.8136 |

TABLE III. DIEBOLD-MARIANO TEST RESULTS (P-VALUES)

| Model Pair | DM Test Statistic | p-value | Conclusion ($\alpha = 0.05$) |
|---|---|---|---|
| GAT-LSTM vs Naïve Bayes | 8.12 | $3.15 \times 10^{-13}$ | Significant (GAT-LSTM is better) |
| GAT-LSTM vs Logistic Regression | 10.02 | $1.12 \times 10^{-17}$ | Significant (GAT-LSTM is better) |
| GAT-LSTM vs Random Forest | 7.55 | $2.50 \times 10^{-11}$ | Significant (GAT-LSTM is better) |
| GAT-LSTM vs XGBoost | 3.54 | $8.50 \times 10^{-04}$ | Significant (GAT-LSTM is better) |

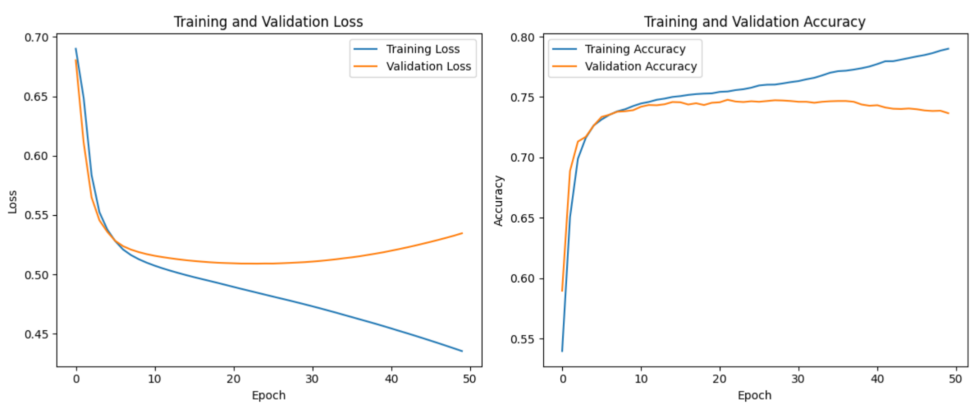


Fig. 6. The cumulative returns of the five major currency pairs during the COVID-19 pandemic.

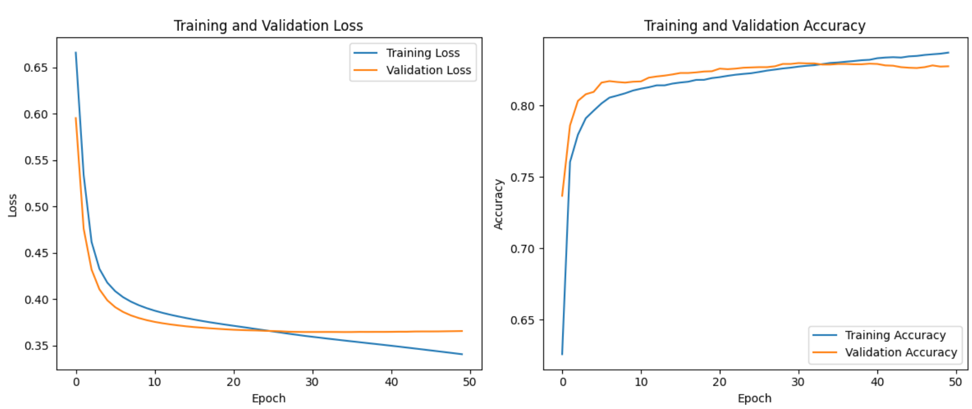


Fig. 7. The cumulative returns of the five major currency pairs during the period test.

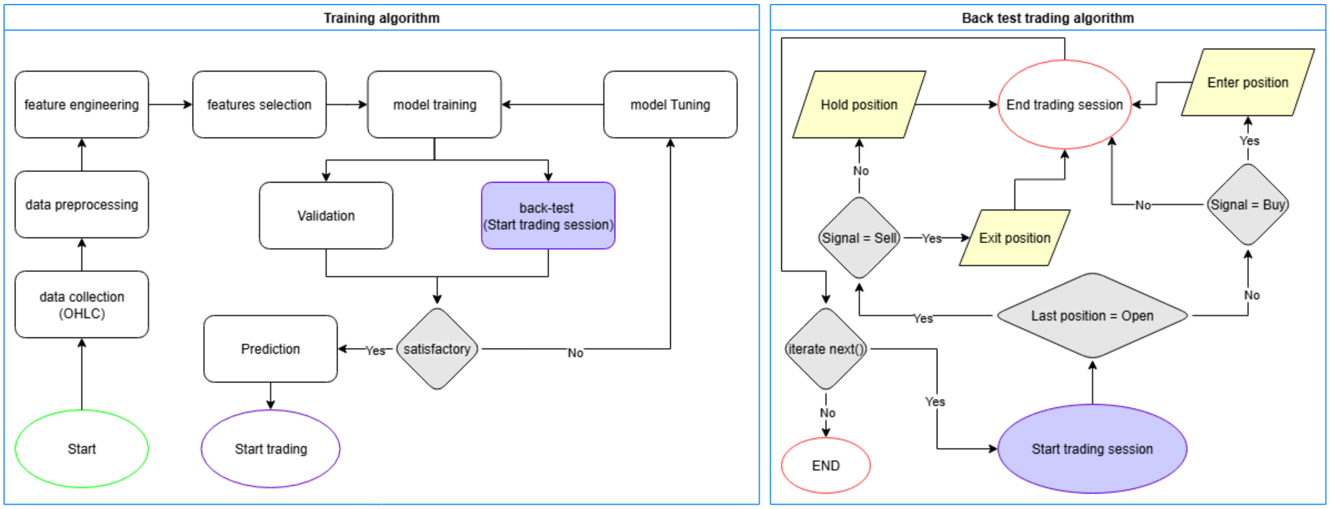


Fig. 8. Our investment strategy and the back-test flowchart.

lating the return profits of forecasted trends. Fig. 8 illustrates our trading strategy and the adopted back-test algorithm.

As illustrated in Fig. 6, our system successfully predicted the future direction of the prices, achieving an accuracy of 83.7% with a total profit of 79% compared to the ideal.

*2) Performance during conflict:* During Global economic crises, the Forex market can experience sharp swings caused by investor panic. The rising index of fear is leading to significant adverse ripple effects within the economy, resulting in a marked increase in market volatility. Every speculative attack on the Forex market frequently coincides with a crisis. We back-tested our model during the COVID-19 pandemic from March 2020 to May 2023.

As shown in the Fig. 7, which shows the cumulative returns of the five major currency pairs during the crisis period, our approach successfully forecasted the future direction of the prices, achieving an average accuracy of 81% corresponding to 73% of the total return profits compared to the ideal. Our suggested trading system can overcome obstacles during times of crisis thanks to the GAT component that can capture dependencies between stocks and currency pairs and anticipate global changes.

*3) Statistical validation of hybrid model superiority:* We used the Diebold-Mariano (DM) test to perform a rigorous statistical comparison of our hybrid GAT-LSTM model against traditional baseline models, assessing their predictive accuracy using a zero-one loss function. The DM test results, detailed in Table III, yielded p-values that were all significantly below the 0.05 significance level ($\alpha$=0.05). This finding provides strong evidence that the performance gains of the GRU-CNN model are statistically significant, meaning that the improved accuracy is due to the model's superior ability to capture complex patterns in the financial time series data, rather than being a result of random chance. This analysis validates our choice of a deep learning architecture for this prediction task.

## V. Conclusion

GAT-LSTM architecture integrates two learning mechanisms simultaneously, spatial and temporal, addressing the complexity and high fluctuations caused by macro-economic, political factors, and news events, etc., offering an effective solution for forecasting Forex prices. As a result, GAT-LSTM is well-equipped to adapt to the rapidly changing nature of Forex markets, providing traders with valuable insights for decision-making. However, currency prices do not depend on other currencies and stock markets only; many other factors can affect the exchange rate. We created and evaluated this solution for a typical period. It could stop being beneficial and lucrative in the future. Additional research must be done to enhance the graph structure, feature engineering, and model architecture for improved performance. Integrating temporal dynamics and inter-currency interactions improves the accuracy and robustness of forecasting models.

## Data Availability Statement

Due to privacy and ethical restrictions, the data that support the findings of this study cannot be made publicly available. However, a simplified version of the codebase is available on request from the corresponding author.

## References

[1] S. Ahmed, S.-U. Hassan, N. R. Aljohani, and R. Nawaz, "Flf-lstm: A novel prediction system using forex loss function," *Applied Soft Computing*, vol. 97, p. 106780, 2020.

[2] R. D. Huang and R. W. Masulis, "Fx spreads and dealer competition across the 24-hour trading day," *The Review of Financial Studies*, vol. 12, p. 61–93, 1999.

[3] D. C. Yıldırım, I. H. Toroslu, and U. Fiore, "Forecasting directional movement of forex data using lstm with technical and macroeconomic indicators," *Financial innovation*, vol. 7, p. 1, 2021.

[4] Y. Bilgin, S. M. Camgoz, M. B. Karan, and Y. Yildiz, *Understanding the investment behavior of individual investors: an empirical study on forex markets*. IGI Global Scientific Publishing, 2021.

[5] L. Puspitawati, "Information system for forex investment and their effects on investment growth in foreign currencies," *International Journal of Research and Applied Technology (INJURATECH)*, 2021.

[6] M. Ozturk, I. H. Toroslu, and G. Fidan, "Heuristic based trading system on forex data using technical indicator rules," *Applied Soft Computing*, vol. 43, p. 170–186, 2016.

[7] M. A. Junior, P. Appiahene, O. Appiah, and C. N. Bombie, "Forex market forecasting using machine learning: Systematic literature review and meta-analysis," *Journal of Big Data*, vol. 10, p. 9, 2023.

[8] H. Law, S. Aghabozorgi, S. Lim, Y. Teh, and T. Herawan, "Processing of kuala lumpur stock exchange resident on hadoop mapreduce," in *Proceedings of the International Conference on Data Science (ICDATA)*. Proceedings of the International Conference on Data Science (ICDATA), 2014, p. 1.

[9] G. McLeod, "Forex market size: A traders advantage," *DailyFX (January 23).*, 2014.

[10] Y. Pei, J. Cartlidge, A. Mandal, D. Gold, E. Marcilio, and R. Mazzon, "Cross-modal temporal fusion for financial market forecasting," *arXiv preprint arXiv:2504.13522*, 2025.

[11] L. Abednego, "Forex trading robot with technical and fundamental analysis," *Journal of Computers*, p. 1089–1097, 2018.

[12] K. Cao and H. You, "Fundamental analysis via machine learning," *Financial Analysts Journal*, vol. 80, p. 74–98, 2024.

[13] L. Abednego, C. E. Nugraheni, and A. Fedora, "Forex sentiment analysis with python," *International Journal of Advanced Research in Economics and Finance*, vol. 4, p. 46–55, 2022.

[14] N. MABROUK, M. CHIHAB, Z. HACHKAR, and Y. CHIHAB, "Intraday trading strategy based on gated recurrent unit and convolutional neural network: forecasting daily price direction," *International Journal of Advanced Computer Science and Applications*, vol. 13, 2022.

[15] E. Gerlein, M. Mcginnity, A. Belatreche, and S. Coleman, "Evaluating machine learning classification for financial trading: An empirical approach," *Expert Systems with Applications*, vol. 54, pp. 193–207, Jul. 2016.

[16] C. J. Neely and P. A. Weller, "Technical analysis in the foreign exchange market," *Handbook of exchange rates*, pp. 343–373, 2012.

[17] V. Drakopoulou, "A review of fundamental and technical stock analysis techniques," *Journal of Stock & Forex Trading*, vol. 5, 2016.

[18] T. N. T. Thu and V. D. Xuan, "Using support vector machine in forex predicting," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*. 2018 IEEE International Conference on Innovative Research and Development (ICIRD), 2018, p. 1–5.

[19] L. Tang, H. Pan, and Y. Yao, "K-nearest neighbor regression with principal component analysis for financial time series prediction," in *Proceedings of the 2018 international conference on computing and artificial intelligence*. Proceedings of the 2018 international conference on computing and artificial intelligence, 2018, p. 127–131.

[20] S. Zhelev and D. R. Avresky, "Using lstm neural network for time series predictions in financial markets," in *2019 IEEE 18th international symposium on network computing and applications (NCA)*. 2019 IEEE 18th international symposium on network computing and applications (NCA), 2019, p. 1–5.

[21] L. Dongdong, S. Yuan, M. Li, and Y. Xiang, "An empirical study of machine learning algorithms for stock daily trading strategy," *Mathematical Problems in Engineering*, vol. 2019, pp. 1–30, Apr. 2019.

[22] D. Lv, Z. Huang, M. Li, and Y. Xiang, "Selection of the optimal trading model for stock investment in different industries," *PLOS ONE*, vol. 14, p. e0212137, Feb. 2019.

[23] B. Yue, J. Rao, and Y, "A deep learning framework for financial time series using stacked autoencoders and long short term memory," *PLoS ONE*, vol. 12, p. 1–24, 2017.

[24] R. Akita, A. Yoshihara, T. Matsubara, and K. Uehara, *Deep learning for stock prediction using numerical and textual information*. Okayama, Japan: IEEE, Jun. 2016.

[25] Z. Hu, Y. Zhao, and M. Khushi, "A survey of forex and stock price prediction using deep learning," *Applied System Innovation*, vol. 4, p. 9, Feb. 2021.

[26] H. Jia, "Deep learning algorithm-based financial prediction models," *Complexity*, vol. 2021, p. 5560886, 2021.

[27] S. Galeshchuk, "Neural networks performance in exchange rate prediction," *Neurocomputing*, vol. 172, p. 446–452, 2016.

[28] Z. Hu, W. Liu, J. Bian, X. Liu, and T.-Y. Liu, "Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction," in *Proceedings of the eleventh ACM international conference on web search and data mining*. Proceedings of the eleventh ACM international conference on web search and data mining, 2018, p. 261–269.

[29] S. Galeshchuk and S. Mukherjee, "Deep networks for predicting direction of change in foreign exchange rates," *Intelligent Systems in Accounting, Finance and Management*, vol. 24, p. 100–110, 2017.

[30] M. Fisichella and F. Garolla, "Can deep learning improve technical analysis of forex data to predict future price movements?" *Ieee Access*, vol. 9, pp. 153 083–153 101, 2021.

[31] J. Wang, "Multivariate time series forecasting and classification via gnn and transformer models," *Journal of Computer Technology and Software*, vol. 3, 2024.

[32] Y. Wang and T. Aste, "Network filtering of spatial-temporal gnn for multivariate time-series prediction," in *Proceedings of the Third ACM International Conference on AI in Finance*. Proceedings of the Third ACM International Conference on AI in Finance, 2022, p. 463–470.

[33] S. Iyer and T. T. Hou, "Gat-gan: A graph-attention-based time-series generative adversarial network," *arXiv preprint arXiv:2306.01999*, 2023.

[34] Y. Liu, H. Zheng, and J. Zhao, "Enhanced water quality prediction by lstm and graph attention network (l-gat): An analytical study of the pearl river basin," *Water Research X*, p. 100383, 2025.

[35] F. H. O. Duarte, G. J. Moreira, E. J. Luz, L. B. Santos, and V. L. Freitas, "Time series forecasting of covid-19 cases in brazil with gnn and mobility networks," in *Brazilian Conference on Intelligent Systems*. Springer, 2023, pp. 361–375.

[36] Y. Bai, G. Li, T. Lu, Y. Wu, W. Zhang, and Y. Feng, "Map matching based on seq2seq with topology information," *Applied Sciences*, vol. 13, p. 12920, 2023.

[37] A. G. Vrahatis, K. Lazaros, and S. Kotsiantis, "Graph attention networks: a comprehensive review of methods and applications," *Future Internet*, vol. 16, no. 9, p. 318, 2024.

[38] K. K. L. Wong, "The structure of neural network," 2024.

[39] F. M. Salem, *Recurrent neural networks*. Springer, 2022.

[40] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent neural networks: A comprehensive review of architectures, variants, and applications," *Information*, vol. 15, p. 517, 2024.

[41] hochreiter Sepp and schmidhuber jürgen, "Memory, long short-term," *Neural Computation*, vol. 9, p. 1735, 1997.

[42] S. Hochreiter and J. Schmidhuber, "Lstm can solve hard long time lag problems," *Advances in neural information processing systems*, vol. 9, 1996.

[43] F. A. Gers and E. Schmidhuber, "Lstm recurrent networks learn simple context-free and context-sensitive languages," *IEEE transactions on neural networks*, vol. 12, p. 1333–1340, 2001.

[44] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," *Neural computation*, vol. 12, p. 2451–2471, 2000.