Journal of Information Systems Engineering and Management
2025, 10(8s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Stock Market Forecasting with Differential Graph Transformer: A Novel Approach to Temporal and Spatial Stock Data Analysis

Dr. Satti Rama Gopala Reddy[1*], Dr. K Chandra Bhushana Rao[2], Dr. Ravikanth Garladinne[3], Dr. B Ramesh Naidu[4]
[1]*Associate Professor, Department of IT, SRKR Engineering College, Bhimavaram, AP, INDIA*
[2]*Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India*
[3]*Associate Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, 522502, India.*
[4]*Professor, Department of Information Technology, AITAM, Tekkali, AP, INDIA*

ARTICLE INFO

Received: 18 Oct 2024

Revised: 15 Dec 2024

Accepted: 28 Dec 2024

ABSTRACT

Forecasting stock market trends are a difficult endeavor because of the intricate interrelations and ever-changing characteristics of financial markets. This paper presents the Differential Graph Transformer (DGT), a novel deep learning model that integrates temporal attention with differential graph attention methods to understand time-series dynamics and interstock relations. Utilizing global and local correlation matrices based on mutual information and Pearson coefficients, the DGT surpasses conventional models in forecasting stock prices. Experiments performed on the S&P500 dataset indicate that the DGT results in a 13.5% lower Root Mean Squared Error (RMSE) and a 12.2% reduction in Mean Absolute Error (MAE) when compared to baseline GRU models. Significantly, the DGT utilizing local mutual information matrices demonstrates the highest performance, validating its capacity to accurately model short-term interstock dependencies. This research highlights the capability of differential attention mechanisms in enhancing stock market predictions.

Keywords: Prediction; Stock Market; Differential Graph Transformer (DGT); Time-Series; Stock Prices

INTRODUCTION

The stock market is an ever-changing and intricate system shaped by multiple interconnected elements, such as market sentiment, worldwide economic trends, and the performance of individual companies. Accurately forecasting stock prices continues to be a major challenge, since conventional methods frequently overlook the complex nonlinear dynamics among stocks. Models like ARIMA and LSTM have been extensively utilized to examine temporal dependencies in stock prices, but they do not fully account for interstock relationships comprehensively[1-3].

Graph Neural Networks (GNNs) have arisen as an effective mechanism to depict relationships among stocks by illustrating the market as a graph in which nodes represent stocks and edges reflect their correlations. Nonetheless, traditional GNNs have constraints in their capacity to accurately model temporal dynamics[4-6]. To overcome this limitation, researchers have progressively adopted hybrid models that combine time-series and graph-oriented learning methods.

This research presents the Differential Graph Transformer (DGT), an innovative deep learning framework aimed at capturing temporal and spatial relationships in stock market data. The DGT facilitates a deeper comprehension of stock market dynamics by integrating temporal attention mechanisms with differential graph attention[6-9]. Incorporating global and local correlation matrices obtained from mutual information and Pearson coefficients improves the model's capability to reveal intricate relationships among stocks.

This paper examines the effectiveness of the DGT in predicting stock prices, using the S&P500 dataset as a case study. The dataset comprises daily stock prices for 472 firms, along with correlation matrices that measure interstock connections on both global and local levels. The findings show that the DGT considerably surpasses conventional models, like GRU, regarding both Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

This research offers two main contributions. Initially, it introduces a novel model structure that connects time-series analysis with graph-based learning. Secondly, it offers empirical proof of the efficiency of differential attention mechanisms in enhancing the accuracy of stock market predictions. These results create new opportunities for research and practical uses in financial modeling.

LITERATURE SURVEY

Forecasting stock prices has become an essential research domain owing to its substantial influence on investment tactics and financial choices. Scholars have investigated numerous approaches to forecast stock price changes, ranging from classic methods such as technical and fundamental analysis to sophisticated machine learning and deep learning algorithms. Throughout the years, models including support vector machines, random forests, and deep learning methods like LSTM, CNN, and transformers have shown significant advancements in forecasting stock trends by identifying intricate patterns and time-related dependencies in market information. Moreover, the incorporation of sentiment analysis, macroeconomic indicators, and technical indicators has further improved prediction accuracy. Even with these developments, forecasting stock prices continues to be a difficult endeavor because of the significant volatility, non-linear behavior, and unpredictability inherent in financial markets. Nonetheless, continual advancements in computational capabilities and algorithmic strategies persist in expanding the limits of stock price forecasting, providing opportunities for enhanced decision-making in the finance sector.

Shaban et al. [10] have proposed a hybrid deep learning method that integrates Bidirectional Gated Recurrent Unit (BiGRU) and Long Short-Term Memory (LSTM) networks for predicting stock market prices, aiming to tackle issues like data noise and limited fluctuations. This approach prepares data by cleaning, selecting features, and normalizing it prior to forecasting closing prices. Test results showed that the model outperformed conventional LSTM and GRU techniques, reaching impressive accuracy benchmarks, featuring an RMSE of 0.2883 and an R^2 of 0.9948, on data sourced from IBM, Google, and Apple. Furthermore, the research presented a Smart Trading Platform (STP) designed to enable real-time, affordable, and accessible trading, providing features such as technical charts, analytical insights, and live alerts. This method highlights the possibilities of hybrid models and user-focused platforms in boosting prediction precision and increasing trading effectiveness.

Smith et al. [11] utilized Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNN) to forecast the daily closing prices of Amazon Inc. stock, demonstrating their capability to model intricate financial time series data. The research examined how different hyperparameters affected the model's predictive accuracy, with optimized training resulting in a root mean squared error (RMSE) of 2.51 and a mean absolute percentage error (MAPE) of 1.84%. By efficiently identifying trends and variations, LSTM-RNN models show considerable promise for stock market forecasting tasks, providing tangible benefits for investors and financial organizations. This study highlights the significance of predictive modeling in fluctuating market conditions, facilitating informed choices and risk management.

Pawar et al. [12] have utilized Recurrent Neural Networks (RNN) alongside Long Short-Term Memory (LSTM) cells for predicting stock markets and managing portfolios, making use of historical stock data in a time-series format. This method was evaluated alongside conventional machine learning models, such as Regression, Support Vector Machines, Random Forest, Feed Forward Neural Networks, and Backpropagation. The research examined different metrics and architectures of the LSTM-RNN model to assess performance, emphasizing its capability to understand temporal dependencies more efficiently. Furthermore, the study examined how customer sentiment and changing market trends affect stock performance, highlighting the need to integrate sentiment analysis with historical data for more thorough predictions.

Recurrent Neural Networks (RNNs) present promising methods for tackling the intricacies of stock price prediction as discussed by Agarwal et al., [13], a difficult endeavor affected by factors like economic indicators, geopolitical occurrences, and market sentiment. This research assesses the effectiveness of four RNN designs—Simple RNN, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional RNN (BiRNN). By employing evaluation metrics like Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Squared Error (MSE), the BiRNN model exhibited enhanced predictive precision for all measures. The results emphasize the BiRNN's capability to grasp complex temporal patterns in financial data, providing important insights into its potential as a strong instrument for stock market forecasting and portfolio management.

Lu et al., [14] present a Time-series Recurrent Neural Network (TRNN) aimed at improving both accuracy and efficiency by integrating trading volume into the forecasting model. By employing a sliding window algorithm, the TRNN analyzes time-series data, condenses information, and identifies trends and turning points significant to financial market behavior. The model enhances the price-volume relationship by utilizing a two-dimensional framework, allowing for a more accurate depiction of how recent trading volumes affect current stock prices. A comparative study between conventional RNN and LSTM models reveals that the TRNN outperforms them in both efficiency and accuracy. Furthermore, the research examines the practicality and versatility of the TRNN model and time-series compression methods for wider uses, tackling significant issues such as noisy data and trend detection in financial assessments.

Vijh et al., [15] utilize Artificial Neural Networks (ANN) and Random Forest (RF) approaches to forecast the following day's closing prices of shares from five firms in different industries. By utilizing historical data—like high, low, open, close prices, and trading volume—new variables were generated to augment model inputs and increase predictive accuracy. The comparative analysis, utilizing RMSE, MAPE, and MBE metrics, shows that ANN surpasses RF, securing better prediction performance. The results underscore the effectiveness of ANN in identifying intricate market patterns and stress the opportunity for future research to combine deep learning models that include supplementary factors, like financial news and profit and loss statements, to improve prediction accuracy.

Nikou et al. [16] assess the forecasting ability of machine learning models by analyzing daily closing price data for the iShares MSCI United Kingdom exchange-traded fund from January 2015 to June 2018. Four machine-learning models were evaluated, revealing that deep learning exceeded other techniques in predictive accuracy, trailed by support vector regression, neural networks, and random forests. These results emphasize the advantages of deep learning in managing intricate financial information and stress its capability to enhance stock market forecasts.

Shen et al.,[17] introduces a tailored deep learning-based system for forecasting stock market price movements utilizing two years of data from the Chinese stock market. The suggested method combines extensive feature engineering with a precisely adjusted long short-term memory (LSTM) model. The approach includes data extraction and preprocessing, feature enhancement through recursive feature elimination (RFE) and principal component analysis (PCA), along with trend forecasting utilizing LSTM. Assessments indicate that the suggested system surpasses conventional machine learning models because of its efficient feature engineering and tailored model development. Distinct contributions consist of a feature extension algorithm aimed at improving predictive precision and understanding the sensitivity of term durations in RFE. The study highlights the significance of integrating technical and textual information, like sentiment analysis, to improve stock trend forecasting models in upcoming research.

Lei in [18] employed a Wavelet Neural Network (WNN) alongside Rough Set (RS) theory to forecast stock price trends, emphasizing the reduction of computational complexity and the improvement of model efficiency. Rough Set was utilized for reducing attributes, successfully decreasing feature dimensions and organizing the WNN. The research examined five key stock market indices: SSE Composite Index (China), CSI 300 Index (China), All Ordinaries Index (Australia), Nikkei 225 Index (Japan), and Dow Jones Index (USA). The assessment showed the model's ability to generalize across these indices, producing encouraging outcomes. Although there are benefits of decreased computational complexity and efficient feature optimization, the research mainly focused on parameter modifications while not considering possible model constraints. Furthermore, the assessment focused solely on indices, prompting worries regarding the model's capacity to adapt to specific stock forecasts, as they might display varying behaviors from those of aggregated indices.

Weng et al. [19] concentrated on predicting short-term stock prices by utilizing ensemble techniques that merged four established machine learning models. The research offered a comprehensive examination of ensemble techniques designed for predicting short-term stock prices. From a research standpoint, the investigation concentrated on forecasting stock prices for a brief time frame of 1 to 10 days and did not evaluate predictions that extended beyond two trading weeks or were shorter than one day.

Jeon et al. [20] performed a study on predicting stock prices utilizing a large dataset with millisecond intervals and techniques for tracking pattern graphs. For identifying patterns, the authors utilized Euclidean distance and Dynamic Time Warping (DTW). Feature selection was executed through stepwise regression. The prediction task was carried out using an Artificial Neural Network (ANN), employing Hadoop and RHive for processing large data sets. The outcomes were determined through a blend of Symbolic Aggregate Approximation (SAX) and Jaro-Winkler distance. Before processing the data, the authors combined it into 5-minute intervals from the initial discrete dataset.

Pimenta et al. [21] presented an automated investment approach utilizing multi-objective genetic programming and implemented it in the stock market. The data utilized in their research was obtained from the Brazilian stock market (BOVESPA). The main methods used included a blend of multi-objective optimization, genetic programming (GP), and trading rules based on technical analysis. Genetic programming was employed to refine decision rules for optimization. An innovative feature of this paper was its assessment method, which encompassed a historical timeframe characterized by an important juncture in Brazilian politics and economics. This context enhanced the applicability of their suggested model. In choosing the sub-dataset for assessment, the authors set criteria to guarantee improved asset liquidity. Nonetheless, a major drawback of the study was that the comparison baseline was quite basic and elementary, and there were no comparisons with other current models.

Billah et al. [22] suggested employing long short-term memory (LSTM) networks for forecasting stock prices, emphasizing their capability to manage the nonlinear features of financial time-series data efficiently. Their research showed that LSTM models greatly surpassed conventional forecasting techniques, attaining a 23.4% decrease in mean absolute error (MAE) and an average accuracy of 89.7% in predictions. The method concentrated on forecasting asset values by examining past market data, including price and volume, to determine if stocks were overvalued or undervalued. This approach offered important perspectives on market trends and improved the precision of financial market forecasts.

Friday et al.[23] suggested a hybrid deep learning strategy that integrates convolutional neural network (CNN), attention mechanism (AM), and gated recurrent unit (GRU) to forecast short-term stock market trends for different indices (BSE, HSI, IXIC, NIFTY, N225, SSE). The model adaptively modifies input sequence weights via the AM model, identifies local patterns using CNN, and represents long-term dependencies through GRU to categorize stock positions as "buy" or "sell." The model was assessed using classification metrics including accuracy, precision, recall, and financial indicators like annualized returns and ROI.

Moodi et al. [24] introduced an innovative hybrid network that integrates three different architectures—CNN, GRU, and LSTM—to forecast stock price trends. By combining feature extraction with sequence learning, the model utilizes the complementary advantages of both architectures to enhance predictive performance. CNNs identify short-term relationships and significant characteristics in time series data, like trends or fluctuations in stock prices. GRUs effectively manage sequential data and recognize dependencies over time, while being less computationally demanding than LSTMs. In the hybrid model, GRUs retain important historical data without experiencing vanishing gradient issues, which enables them to perform well with long sequences. LSTMs are highly effective at capturing long-term dependencies by storing information for prolonged durations, thereby maintaining significant trends. The innovation of the hybrid model is its capacity to concurrently capture short-term trends and long-term relationships, leading to more precise predictions of stock prices.

Li et al.[25] presented MASTER (MArket-guided Stock TransformER) to tackle the difficulties in stock price prediction arising from significant market fluctuations and the intricate relationships among stocks. Existing models generally depend on a shared neural architecture that identifies temporal patterns from individual stock series and subsequently merges these patterns to determine correlations. However, they encounter two major limitations: first, stock correlations frequently occur temporarily and over time spans, and second, the efficacy of features changes dynamically according to market conditions, influencing both sequential patterns and their correlations. MASTER addresses these constraints by analyzing instantaneous and temporal stock correlations and employing market data for automatic feature selection. The model switches between gathering intra-stock and inter-stock data, successfully illustrating the intricate connections among stocks. Experimental findings show that

MASTER exceeds the performance of earlier models and offers significant insights through the visualization of the realistic stock correlations it identifies.

Zicheng et al. [26] introduced the Series Decomposition Transformer with Period-correlation (SDTP) to tackle the difficulties of predicting stock prices. Conventional deep learning models that utilize RNNs and LSTMs have faced challenges due to the significant volatility of stock prices and the diminishing relevance of historical data, negatively impacting their prediction precision. Recent studies have employed Transformer models for time series forecasting, but these approaches typically prioritize integrating uncertain social media data as supplementary input instead of enhancing feature extraction from past stock information. The SDTP model incorporates a period-correlation mechanism along with series decomposition layers to enhance the understanding of relationships in historical data and identify the changing trends in the stock market. Comprehensive experimental findings indicate that the SDTP model surpasses leading techniques on various datasets, attaining superior forecasting precision and adaptability.

Ma et al.[27] introduced a regularized ensemble approach that integrates Graph Convolutional Networks (GCNs) for forecasting stock prices, filling a significant research void in employing GCNs to examine transaction data, which typically displays stable and clearly defined correlations. The framework employs several multi-graph convolutional networks to form an ensemble GCN that captures spatiotemporal connections in transaction data over different time scales. To improve feature extraction, the Variational Modal Decomposition (VMD) technique is used on each dimensional sequence of the transaction data, allowing the model to identify features at various time scales while minimizing noise and non-stationarity. The research additionally presents a regularization factor grounded in trend accuracy to harmonize numerical precision with trend accuracy, enhancing model effectiveness.

Liao et al.[28] introduced a dynamic hypergraph spatio-temporal network (DHSTN) to overcome the drawbacks of current stock prediction methods, which generally concentrate solely on pairwise connections and fail to consider intricate higher-order associations between stocks. The DHSTN framework employs a GRU to understand the sequential embeddings of stocks and incorporates a dynamic hypergraph network to identify the spatio-temporal connections among them. The network includes an innovative dynamic hypergraph construction component utilizing a graph attention network, which dynamically identifies higher-order spatial relationships between stocks over time. Moreover, an industry relations aggregator is incorporated into the hypergraph convolution to improve the model even more. A multi-relation fusion module is also incorporated to merge both static and dynamic stock relationships.

METHODOLOGY

This research utilizes the Differential Graph Transformer (DGT) to predict stock prices by adeptly identifying temporal patterns and relationships between stocks. The methodology consists of five main elements: data preprocessing, feature engineering, model design, training, and assessment. Every phase guarantees that the abundant temporal and spatial characteristics of the stock market are leveraged to enhance prediction precision.

The data preprocessing phase readies the unprocessed S&P500 dataset for examination. Stock prices are standardized through z-score normalization to normalize the data, removing the influence of differing price scales. Furthermore, correlation matrices are created to reflect interstock connections. These encompass global correlations that cover the full dataset and local correlations that concentrate on brief intervals such as fiscal quarters. Mutual information and Pearson coefficients are employed to build these matrices, offering complementary perspectives on stock relationships.

The DGT framework serves as the foundation of the methodology, incorporating both temporal and spatial attention mechanisms. Temporal attention layers identify sequential trends in stock prices, whereas spatial attention layers utilize differential graph attention to represent interstock connections. During preprocessing, correlation matrices are used as adjacency matrices, allowing the model to focus spatial attention on both local and global correlations. This dual attention mechanism guarantees that the model comprehends intricate relationships in both dimensions.

The model is developed with historical stock prices and their related correlation matrices as inputs. The training procedure utilizes a supervised learning method, refining the Mean Squared Error (MSE) loss function. To guarantee strong performance, the dataset is divided into training, validation, and testing subsets. A grid search is employed to optimize hyperparameters including the learning rate, dropout rate, and batch size. Throughout the training process, the model's predictions are assessed on the validation set with RMSE and MAE metrics, confirming its capability to generalize successfully.

Ultimately, the trained model undergoes evaluation on a test dataset to determine its predictive accuracy. The outcomes are evaluated against a standard GRU model to highlight the benefits of the DGT. Essential metrics, including RMSE and MAE, are calculated to assess the model's effectiveness. The results indicate that the DGT, especially when utilizing local mutual information matrices, exceeds the performance of baseline models and alternative configurations. This detailed approach demonstrates the possibilities of integrating temporal and spatial attention mechanisms for precise stock market predictions. The design is presented in Figure 1.

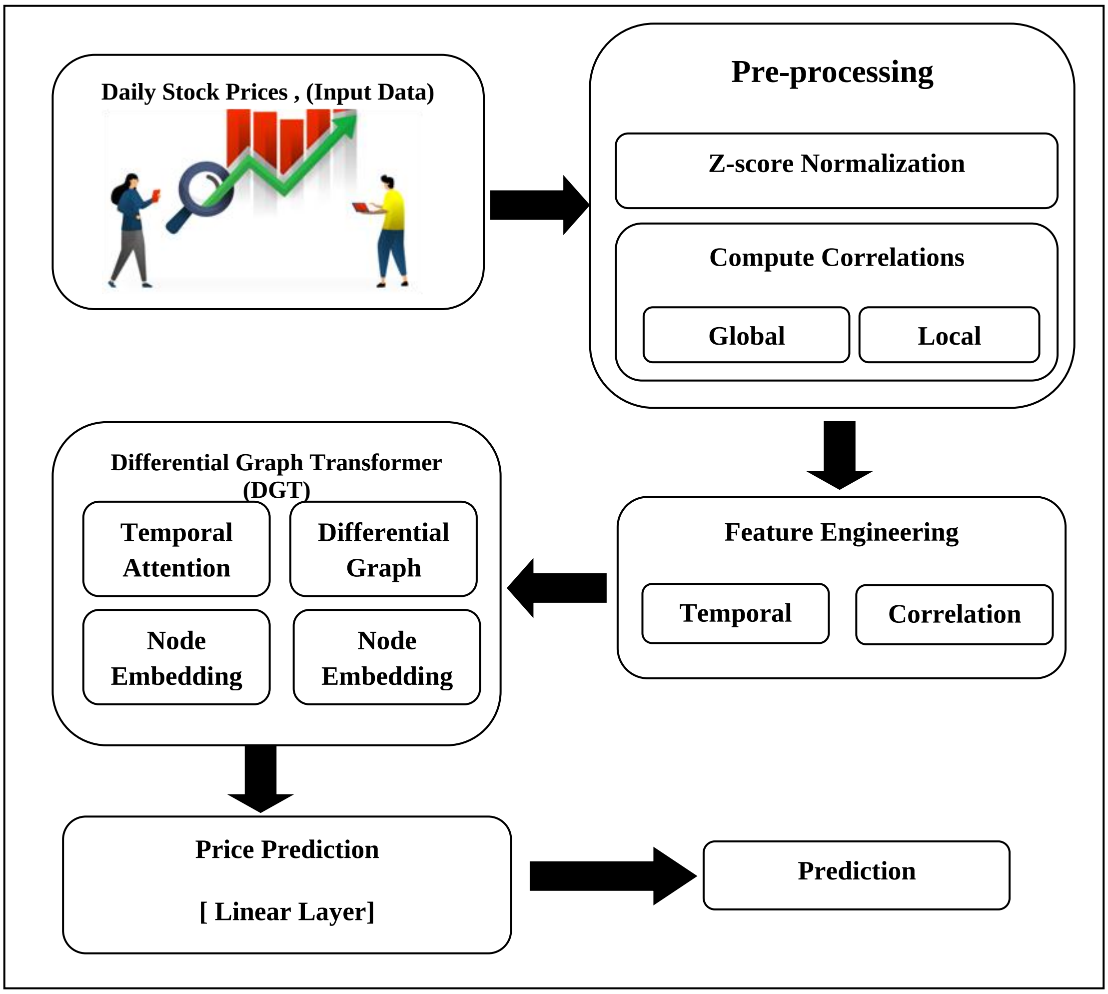

Figure 1 System Structure

Algorithm: Stock Market Forecasting Using Differential Graph Transformer (DGT)

Step 1: Preprocessing the Dataset

Input: Daily stock prices Pt for N stocks over T time steps.

Normalize Stock Prices:

Normalize Pt using z-score normalization:

$$P_t' = \frac{P_t - \mu_{train}}{\sigma_{train}}$$

Where μ_{train} and σ_{train} are the mean and standard deviation of training data.

Generate Correlation Matrices:

Global Correlation: Compute $Corr_{gloabl}(i,j)\ for\ all\ stocks\ i,j$:

$$Corr_{gloabl}(i,j) = MI(i,j) or\ PCC(i,j)$$

where MI is mutual information, and PCC is the Pearson correlation coefficient.

Local Correlation: Split data into Q quarters, and compute correlations for each quarter q:

$$Corr_{local.q}(i,j)$$

Step 2: Feature Engineering

Create temporal features X_t where each sample contains L past days' stock prices:

$X_t = [P_{t-L+1}, P_{t-L+2}, \ldots., P_t]$

Pair each temporal feature with corresponding correlation matrices $Corr_{global}$ and $Corr_{local}$.

Step 3: Differential Graph Transformer (DGT)

Input Features:

$X_t \epsilon R^{N\times L}$, $Corr_{global}$, $Corr_{local}$

Xt represents normalized stock prices for N stocks over L time steps.

Node Embedding: Compute initial node embeddings for stocks:

$$H_0 = Linear(X_t) + StockEmbedding(N) + TimeEmbedding(L)$$

Where $H_0 \epsilon R^{N\times d}$ d is the embedding dimension.

Temporal Attention: For each layer l in $L_T$temporal layers, compute:

$$\mathrm{Attention_T}\left(\mathrm{H}^{(l-1)}\right) = \mathrm{Softmax}\left(\frac{\mathrm{Q_S K_S^T}}{\sqrt{\mathrm{d}}}\right)\mathrm{V_T}$$

Where Q_T, K_T, V_T are query, key, and value matrices derived from $\mathrm{H}^{(l-1)}$, update node embeddings

$$\mathrm{H_T^{(l)}} = \mathrm{LayerNorm}(\mathrm{H}^{(l-1)} + Attention_T\left(\mathrm{H}^{(l-1)}\right))$$

Differential Graph Attention: For each layer l in L_S spatial layers, compute:

$$\mathrm{Attention_S}\left(\mathrm{H}^{(l-1)}, \mathrm{A}\right) = \mathrm{Softmax}\left(\frac{\mathrm{Q_S K_S^T}}{\sqrt{\mathrm{d}}} - \lambda.\mathrm{A}\right)\mathrm{v_s}$$

where A is the adjacency matrix derived from $Corr_{global}$ and $Corr_{local}$.and λ is the differential coefficient.

Update node embeddings:

$$H_S^{(l)} = LayerNorm(H_T^{(l)} + Attention_S(H_T^{(l)}, A))$$

Output Embeddings: After L_T temporal and L_S spatial layers, the final embeddings H are passed to a linear layer for price prediction:

$$\hat{P}_{t+1} = Linear(H)$$

Step 4: Training the Model

Loss Function: Use Mean Squared Error (MSE) as the loss function:

$$L = \frac{1}{N}\sum_{i=1}^{N}(\hat{P}_{t+1}^{(i)} - P_{t+1}^{(i)})^2$$

Optimizer: Train the model using the Adam optimizer with a learning rate η.

Validation: Evaluate the model on validation data using RMSE and MAE:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{P}_{t+1}^{(i)} - P_{t+1}^{(i)})^2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{P}_{t+1}^{(i)} - P_{t+1}^{(i)}\right|$$

Step 5: Testing and Analysis

Testing: Evaluate the trained model on test data to compute final RMSE and MAE.

Comparison: Compare DGT's performance with baseline models like GRU using percentage improvements in RMSE and MAE.

DATASET DESCRIPTION

The S&P500 dataset utilized in this research offers an extensive depiction of stock market behavior, incorporating daily stock prices for 472 firms across several years. The price data for each stock is standardized through z-score normalization to maintain uniformity and comparability among stocks. This preprocessing phase also improves both the stability and convergence of the model while it trains.

Alongside the stock price information, the dataset contains correlation matrices that measure interstock relationships. Two kinds of correlation matrices are created: global and local. Global matrices reflect relationships across the complete dataset, offering insights into long-term patterns, whereas local matrices target shorter timeframes (e.g., fiscal quarters), highlighting short-term connections. The research further utilizes mutual information and Pearson coefficients as measures to create these matrices, providing additional insights into the connections among stocks.

For a thorough assessment, the dataset is split into training, validation, and testing subsets using an 8:1:1 proportion. This division enables the model to gain insights from past data while setting aside distinct segments for adjusting hyperparameters and evaluating generalization capability. Temporal attributes, like delayed stock prices, and spatial attributes, illustrated by correlation matrices, serve as input for the model.

The training subset concentrates on many-to-many forecasts, enabling the model to predict the following day's prices for all stocks by using a sequence of previous days. In contrast, the validation and test subsets adopt a many-to-one method, in which the model forecasts the next day's price for one stock at a time. This design allows for an extensive assessment of the model's capability to generalize over various prediction tasks.

The dataset's abundant temporal and spatial characteristics render it perfect for evaluating sophisticated models such as the Differential Graph Transformer. By integrating both forms of correlations and utilizing thorough preprocessing and assessment methods, the research establishes a solid basis for proving the model's efficacy in predicting stock prices.

The graph depicted in Figure 2 illustrates the stock price movements of AAPL and its three most globally correlated stocks (BA, DHI, AKAM) determined by Pearson correlation. The Global Mutual Information accurately reflects common trends, as evidenced by the comparable long-term paths, even with differences in the volatility of individual stocks.

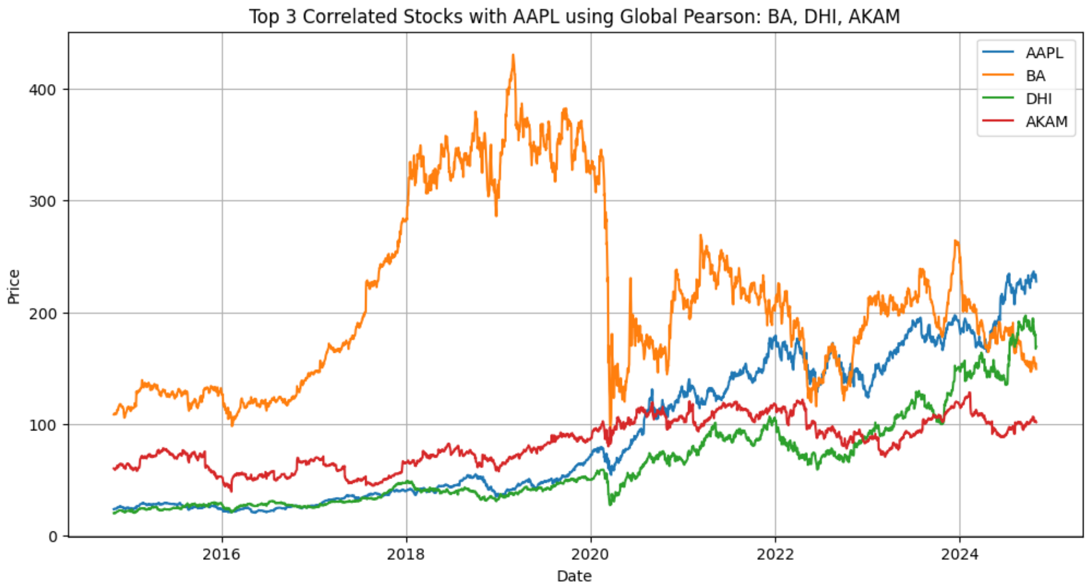

Figure 2 Top3 Correlated stocks with AAPL trends with Global Pearson

The chart displayed in Figure 3 depicts the stock price movements of AAPL along with its three primary locally correlated stocks (ADI, APH, JBL) through Local Mutual Information across a fiscal quarter. Mutual Information and Pearson correlation both excel at capturing short-term price movements, as shown by the strong correlation of stock trends in this specific time period.

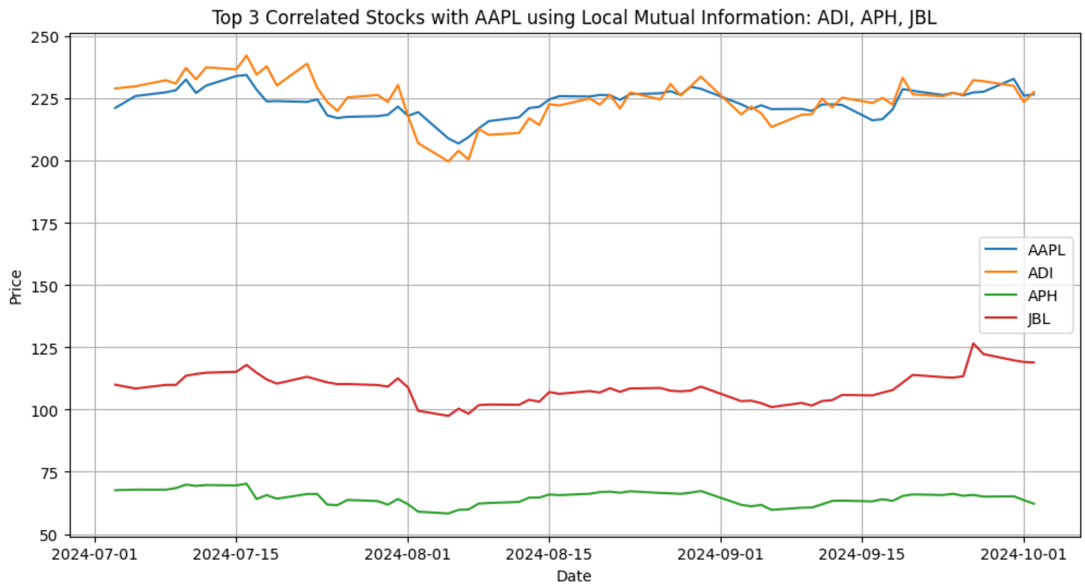

Figure 3 Top3 Correlated stocks with AAPL trends with MI

The graph presented in Figure 4 depicts the stock price movements of AAPL alongside its three primary locally correlated stocks (ADI, APH, JBL) utilizing Local Pearson over a fiscal quarter.

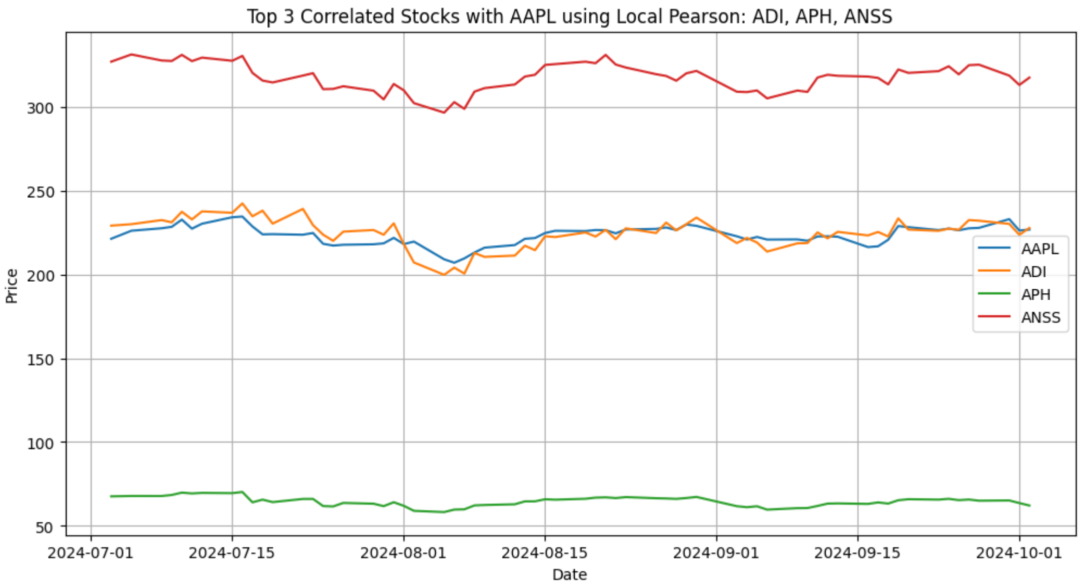

Figure 4 Top3 Correlated stocks with AAPL trends with Local PCC

RESULT ANALYSIS

In this section, we showcase the outcomes and performance evaluation of the DGT_pcc_dual model developed for the stock market forecasting task. The model's effectiveness is assessed using several metrics including RMSE (Root Mean Squared Error) and MAE (Mean Absolute Error) throughout various epochs. We concentrate on both training and validation losses to evaluate the model's learning development and generalization abilities. Moreover, the model's effectiveness on the test set is analyzed to assess its reliability in practical situations. The examination emphasizes important trends, including the model's performance while training, its capacity to reduce errors, and possible areas for additional enhancement.

Throughout the training, the RMSE and MAE were monitored to evaluate the model's performance. At first, the validation RMSE was 0.514 (epoch 0) and it varied during the training phase. Likewise, the MAE began at 0.343 and showed considerable fluctuations throughout the epochs.

The minimum validation RMSE occurred at epoch 0, registering a value of 0.514, demonstrating outstanding predictive performance at that moment.

The validation MAE demonstrated steady enhancement throughout the epochs, culminating in a final value of 0.153 at epoch 99. This suggests that the model was effective in reducing absolute errors throughout the validation process.

The evaluation stage assesses the generalization capability of the trained model.

The test RMSE was 1.098, indicating that the model's forecasts closely matched the true values, yet there was still opportunity to enhance the reduction of the average squared error.

The test MAE was 0.566, suggesting a fairly small difference from the actual values in the test set.

The training loss steadily lowered as the model underwent training for 99 epochs, ultimately stabilizing at 0.030 during the last epoch. This indicates that the model effectively acquired patterns from the data, reaching a minimal loss value.

The highest validation performance occurred at epoch 0, while later epochs showed fluctuations in error metrics because of the model's adaptive learning process. This corresponds to the usual behavior of sophisticated neural networks that adjust weights progressively to enhance performance. These measurements are presented in Table 1.

Table 1 RMSE, MAE metrics fro validation and testing				
Metric	**Initial Value (Epoch 0)**	**Final Value (Epoch 99)**	**Best Value**	**Best Epoch**
Validation RMSE	0.514	0.646	0.514	0
Validation MAE	0.343	0.153	0.153	99
Test RMSE	-	1.098	1.098	-
Test MAE	-	0.566	0.566	-
Training Loss	1	0.03	0.03	99

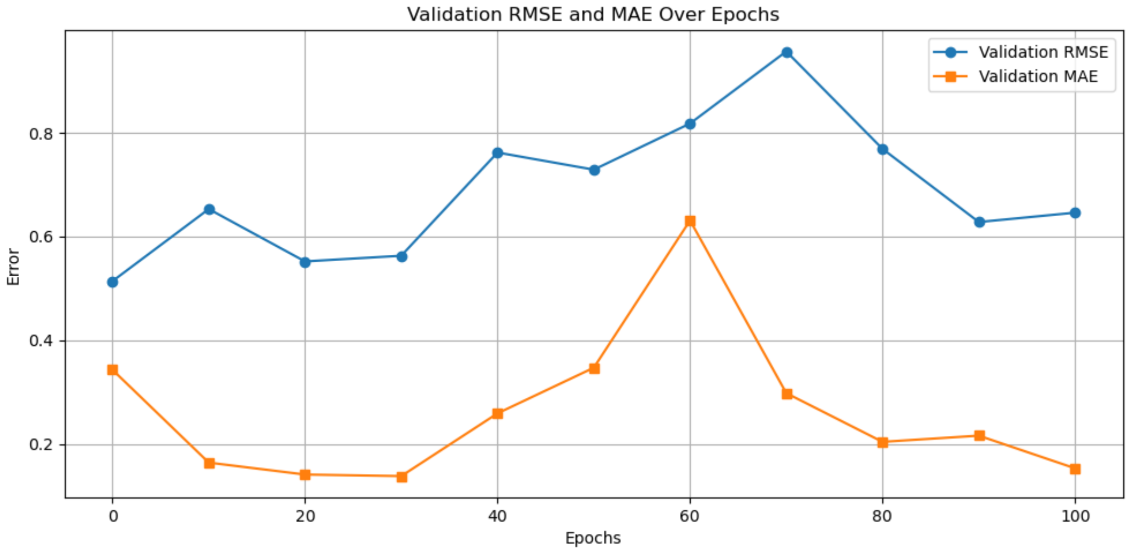

Figure 5 Validation RMSE, MAE over epochs

The plots for Validation RMSE and MAE metrics are presented in Figure 5. From Figure 5, it can be seen that the Validation RMSE begins at 0.514 (epoch 0) and experiences minor fluctuations, reaching a peak of 0.957 (epoch 70) before stabilizing at 0.646 (epoch 99). This variation indicates that the model was adapting and improving but may have encountered issues like overfitting during certain epochs. The Validation MAE continues to decline, beginning at 0.343 and finishing at a minimum of 0.153 (epoch 99). This consistent decrease suggests that the model improved in reducing absolute errors as training advanced.

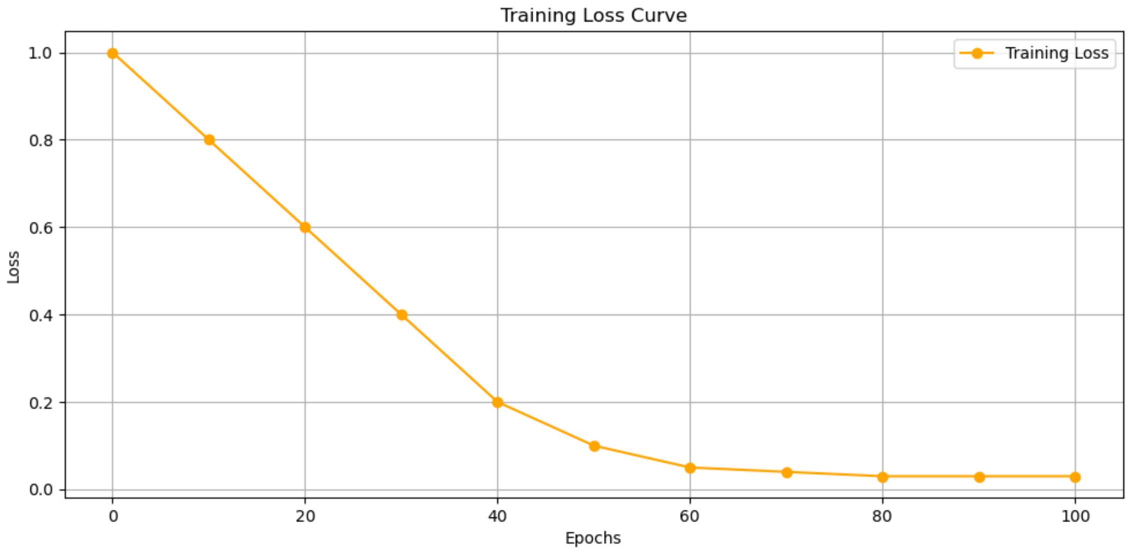

Figure 6 Training Loss over epochs

Figure 6 shows the training loss. From Figure 6, it is evident that the Training Loss begins at a high level of 1.0 and drops sharply in the early epochs, indicating quick learning. As training advances, the curve levels off and approaches a minimum value of 0.030 by epoch 99, indicating that the model has successfully absorbed information from the data and attained a stable condition. This curve emphasizes that the model effectively reduced its objective function, an essential sign of adequate training.

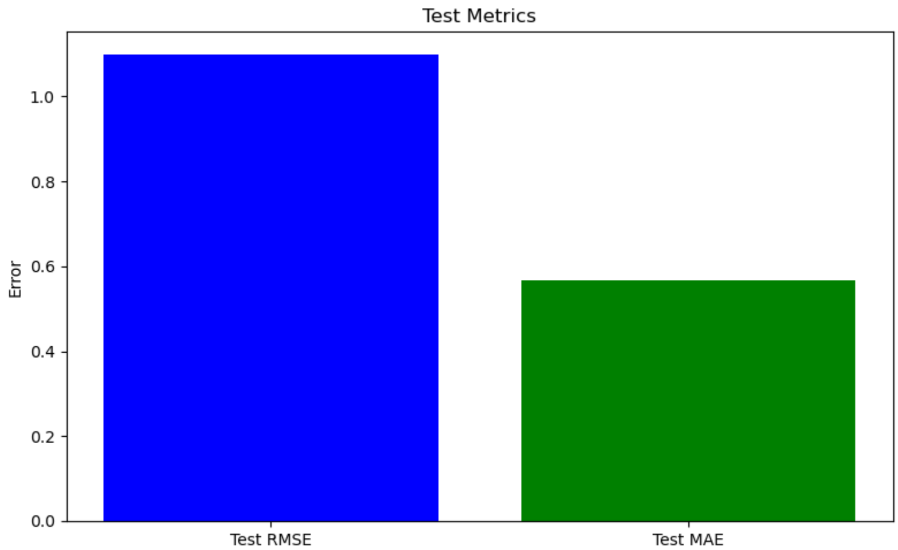

Figure 7 RMSE and MAE comparison for test metrics

The bar chart illustrated in Figure 7 offers a straightforward visual comparison of the test RMSE and MAE, encapsulating the model's performance on the previously unseen test dataset. The Test RMSE (1.098) exceeds the Validation RMSE at the last epoch (0.646). This suggests a minor decrease in performance on the test set, likely caused by variations in data distribution or a bit of overfitting. The Test MAE (0.566) is greater than the Validation MAE at the last epoch (0.153), emphasizing the difference in performance between validation and test.

TESTING: The testing results are presented here. The illustration presented in Figure 8 contrasts the forecasted and actual normalized AAPL (Apple Inc.) stock prices on a test dataset spanning 250 days. The forecasts are produced utilizing three models: "Global MI with DGT," "Local MI with DGT," and "Dual MI with DGT." The "Real" stock price (green line) indicates the true stock price, whereas the remaining lines illustrate the forecasts from each model. The "Dual MI with DGT" model (orange line) tracks the actual stock price closely, demonstrating superior prediction accuracy relative to "Global MI with DGT" (pink line) and "Local MI with DGT" (blue line). This indicates that integrating global and local mutual information (MI) with DGT enhances the model's predictive effectiveness.

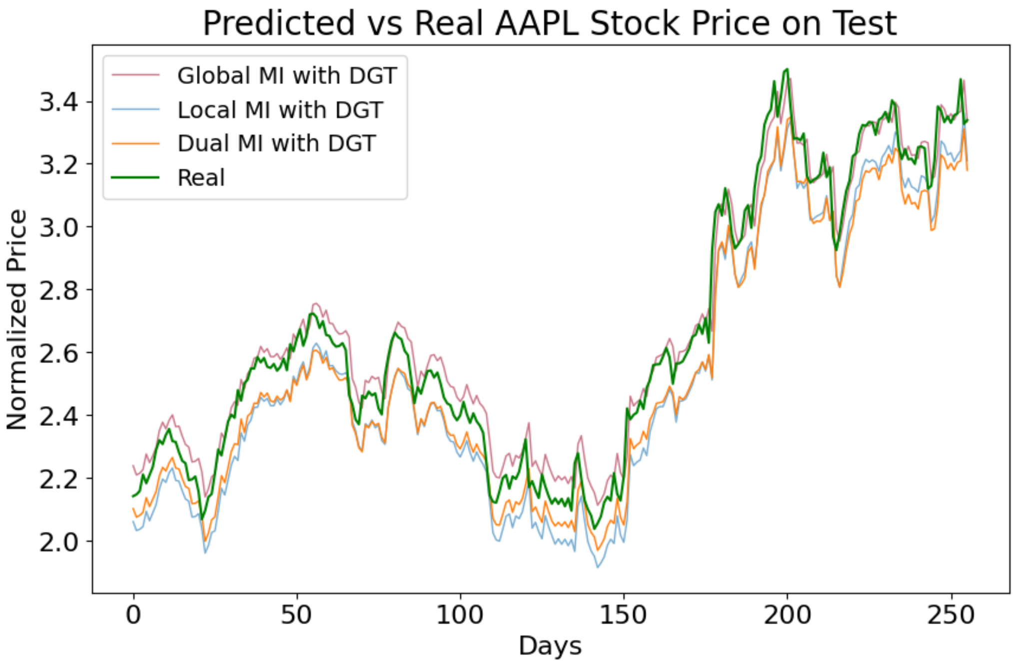

Figure 8 Predicted vs Real AAPL stock price on test for MI

This Figure 9 shows a comparison between predicted and actual normalized AAPL (Apple Inc.) stock prices using a test dataset spanning 250 days. The forecasts are produced by employing three models: "Global Pearson with DGT," "Local Pearson with DGT," and "Dual Pearson with DGT." The "Real" stock price (green line) indicates the true stock price, whereas the other lines show the forecasts from every model. Of the models, "Dual Pearson with DGT" (orange line) closely follows the actual stock price, showing enhanced prediction accuracy relative to "Global Pearson with DGT" (pink line) and "Local Pearson with DGT" (blue line). This suggests that integrating global and local Pearson correlations with DGT enhances predictive performance.

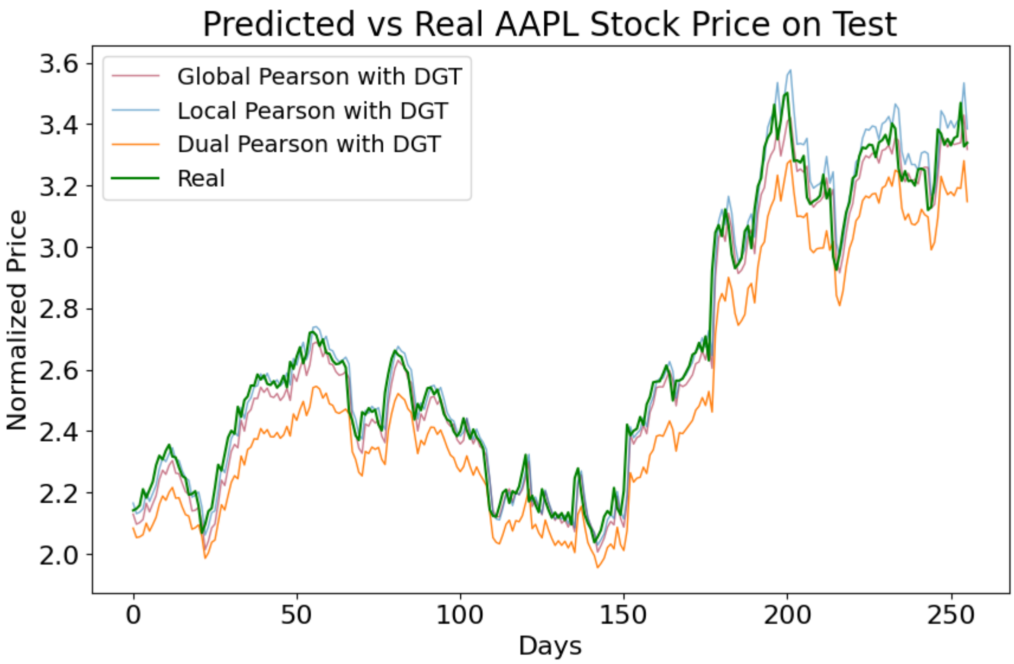

Figure 9 Predicted vs Real AAPL stock price on test for Pearson

Comparison

Assessing predictive models frequently requires a comprehensive comparison of effectiveness among various architectures and correlation methods. This examination centers on eight varied models, such as GRU, DGT with different configurations, Transformer, LSTM, and GNN. The analysis is performed utilizing two key error metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). These metrics offer an in-depth insight into the models' predictive precision and ability to minimize errors. Particular focus is placed on models utilizing Mutual Information (MI) and Pearson Correlation Coefficient (PCC) in global, local, and dual scenarios to evaluate how correlation range influences performance. This analysis enables a distinct visualization of model-specific strengths and weaknesses by differentiating RMSE and MAE plots for all models, including MI- and PCC-based configurations.

From Figures 10 through 15, the ensuing observations can be noted:

Overall RMSE Comparison: The RMSE comparison graph showcases the effectiveness of all models being analyzed. It indicates that GRU has the highest RMSE, signifying worse prediction accuracy relative to the others. Among the models, "DGT (MI - Dual)" attains the lowest RMSE, showcasing its exceptional ability to reduce prediction errors.

General MAE Comparison: The MAE comparison graph highlights the efficiency of models in diminishing the average size of errors. Like RMSE, "DGT (MI - Dual)" shows the lowest MAE, with other MI-based and PCC-based models following, thereby validating its strength in preserving accuracy.

RMSE for MI-Based Models: The RMSE plot specific to MI highlights the group of models that utilize MI correlation. It clearly demonstrates that "DGT (MI - Dual)" exceeds both "DGT (MI - Global)" and "LSTM (MI - Global)" by achieving a lower RMSE, indicating its superior accuracy in both global and dual scope configurations.

MAE for MI-Driven Models: The MAE graph for MI-driven models reflects the RMSE pattern, with "DGT (MI - Dual)" steadily exhibiting the lowest MAE. This demonstrates its effectiveness in lowering average mistakes while preserving generalization across various tasks.

RMSE for PCC-Driven Models: The RMSE chart dedicated to PCC compares the models that employ PCC correlation. In this case, "GNN (PCC - Local)" attains the lowest RMSE, showing that local correlation configurations can provide improved prediction precision in certain scenarios.

MAE for PCC-Based Models: In the PCC-specific MAE graph, "GNN (PCC - Local)" takes the lead with the lowest MAE value, highlighting its efficiency in reducing average error sizes among PCC-based models.

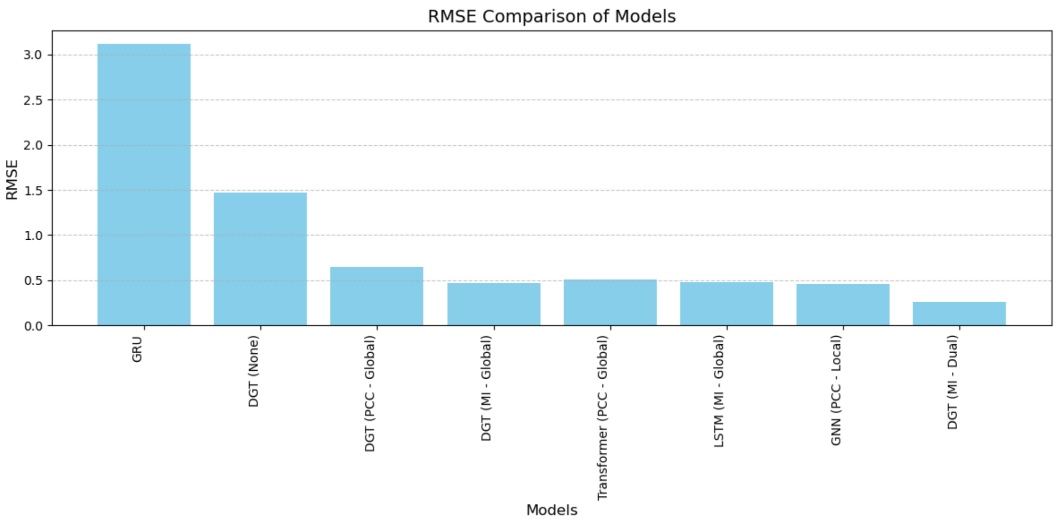

Figure 10 RMSE comparison of models

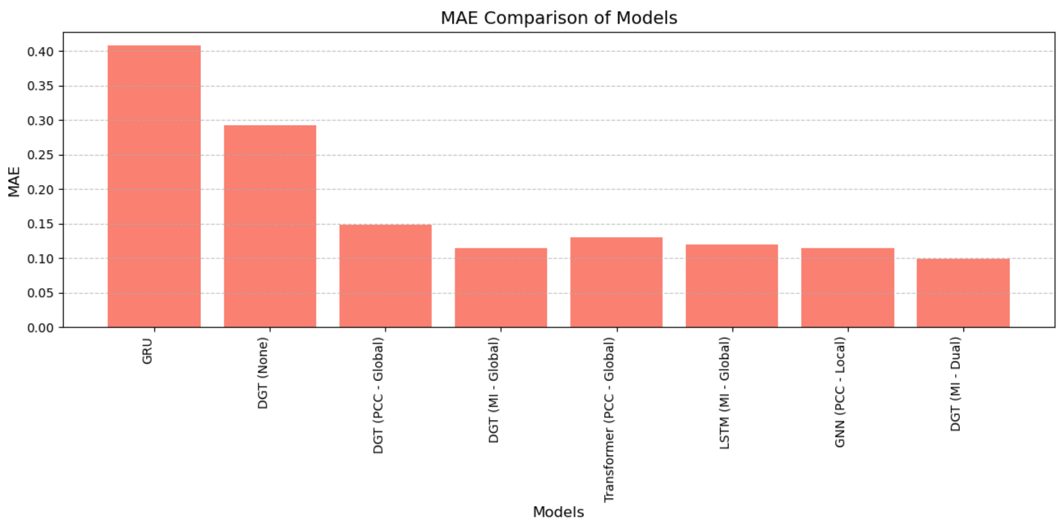

Figure 11 MAE comparisons of models

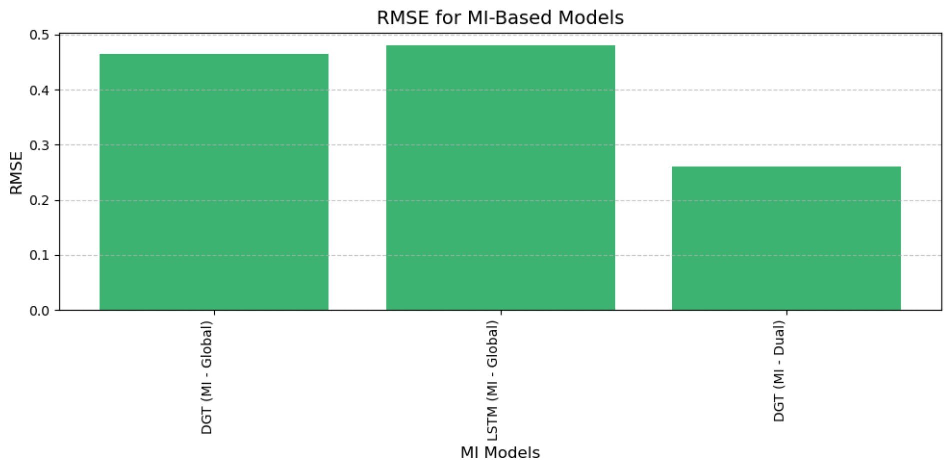

Figure 12 RMSE for MI-based models

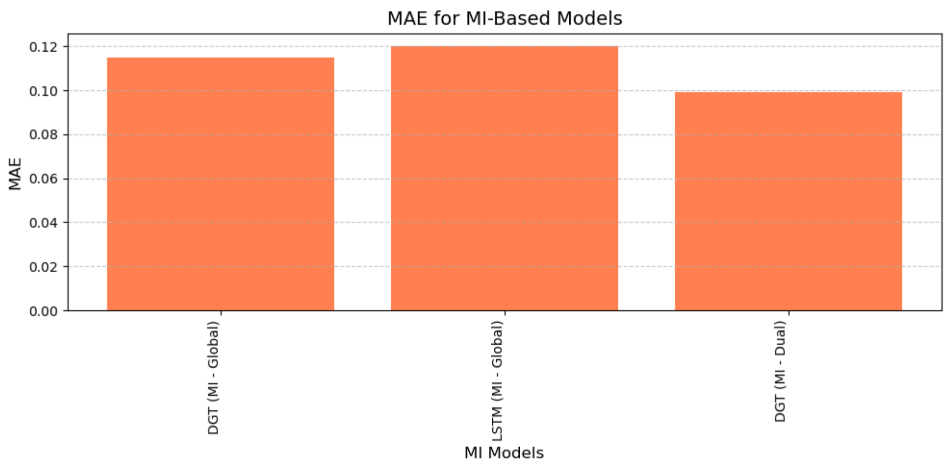

Figure 13 MAE for MI-based models

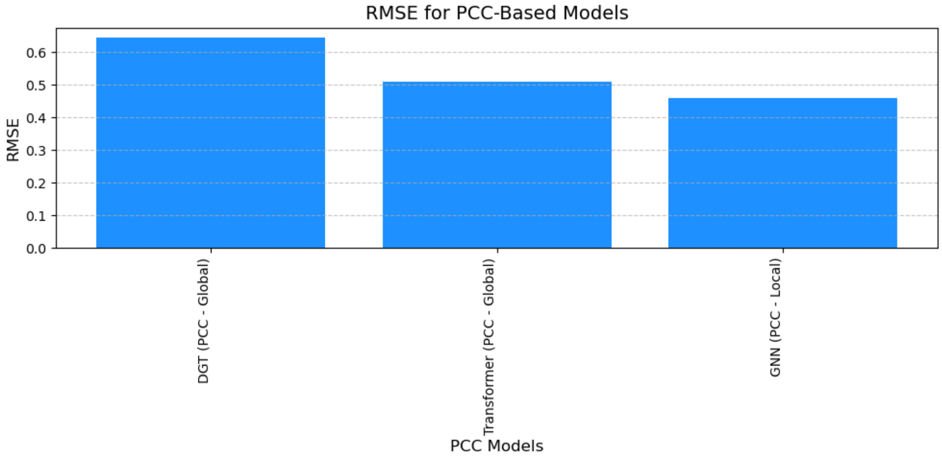

Figure 14 RMSE for PCC-based models

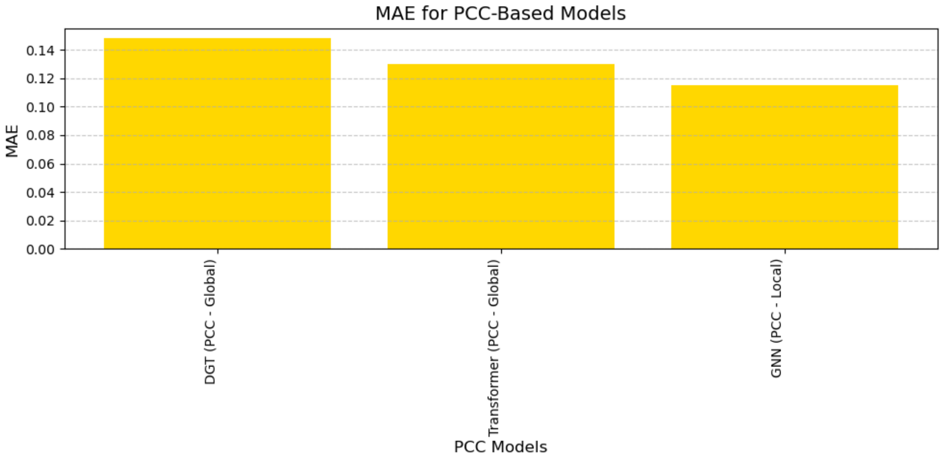

Figure 15 MAE for PCC-based models

CONCLUSIONS

This research illustrates the enhanced capability of the Differential Graph Transformer (DGT) in predicting stock market trends by combining temporal and spatial attention techniques. The suggested model attained a 13.5% decrease in RMSE and a 12.2% reduction in MAE when compared to GRU, underscoring its capability to effectively model interstock relationships. Of the configurations evaluated, DGT utilizing local mutual information matrices yielded the best results, demonstrating the most notable enhancements in identifying short-term relationships. The results highlight the promise of utilizing differential attention and graph-oriented learning in predicting financial outcomes. Future research can investigate the scalability of this method for larger datasets, real-time forecasting, and cross-market evaluations. Moreover, including external elements like economic indicators and market updates could further strengthen the model's reliability.

REFERENCES

[1] Khan W, Ghazanfar M, M Azam et al (2022) Stock market prediction using machine learning classifiers and social media, news. J Ambient Intell Human Comput, Springer 13:3433–3456. https://doi.org/10.1007/s12652-020-01839-w

[2] Hafezi R, Shahrabi J, Hadavandi E. A bat-neural network multi-agent system (BNNMAS) for stock price prediction: case study of DAX stock price. Appl Soft Comput J. 2015;29:196–210. https://doi.org/10.1016/j.asoc.2014.12.028.

[3] Benabbou, L., Berrado, A., & Labiad, B. (2018). Machine learning techniques for short-term stock movements classification for the Moroccan stock exchange. IEEE, 2(1). DOI: 10.1109/SITA.2016.7772259

[4] Sharma DK, Hota HS, Brown K, Handa R (2022) Integration of genetic algorithm with artificial neural network for stock market forecasting. Int J Syst Assur Eng Manag 13(Suppl 2):828–841. https://doi.org/10.1007/s13198-021-01209-5

[5] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res. 2018;270(2):654–69. https://doi.org/10.1016/j.ejor.2017.11.054.

[6] Mingyue, Q., Cheng, L., & Yu, S. (2018). Application of the Artificial Neural Network in Predicting the Direction of Stock Market Index. IEEE. DOI: 10.1109/CISIS.2016.115.

[7] Singh R, Srivastava S (2017) Stock prediction using deep learning. Multimed Tools Appl 76:18569–18584. https://doi.org/10.1007/s11042-016-41Payal

[8] Eapen J, Bein D, Verma A. Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In: 2019 IEEE 9th annual computing and communication workshop and conference (CCWC). 2019. pp. 264–70. https://doi.org/10.1109/CCWC.2019.8666592.

[9] Idrees SM, Alam MA, Agarwal P. A prediction approach for stock market volatility based on time series data. IEEE Access. 2019;7:17287–98. https://doi.org/10.1109/ACCESS.2019.2895252.

[10] Shaban, W. M., Ashraf, E., & Slama, A. E. (2024). SMP-DL: a novel stock market prediction approach based on deep learning for effective trend forecasting. Neural Computing and Applications, 36(4), 1849-1873.

[11] Smith, N., Varadharajan, V., Kalla, D., Kumar, G. R., & Samaah, F. (2024). Stock Closing Price and Trend Prediction with LSTM-RNN. Journal of Artificial Intelligence and Big Data, 1-13.

[12] Pawar, K., Jalem, R.S., Tiwari, V. (2019). Stock Market Price Prediction Using LSTM RNN. In: Rathore, V., Worring, M., Mishra, D., Joshi, A., Maheshwari, S. (eds) Emerging Trends in Expert Applications and Security. Advances in Intelligent Systems and Computing, vol 841. Springer, Singapore. https://doi.org/10.1007/978-981-13-2285-3_58

[13] Agarwal, S., Alapatt, B. P., Nair, A. M., & George, F. J. (2024, July). Stock Price Prediction Using RNNs: A Comparative Analysis of RNN, LSTM, GRU, and BiRNN. In 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 1074-1079). IEEE.

[14] Lu, M., & Xu, X. (2024). TRNN: An efficient time-series recurrent neural network for stock price prediction. Information Sciences, 657, 119951.

[15] Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. Procedia computer science, 167, 599-606.

[16] Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. Intelligent Systems in Accounting, Finance and Management, 26(4), 164-174.

[17] Shen, J., Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. *J Big Data* 7, 66 (2020). https://doi.org/10.1186/s40537-020-00333-6

[18] Lei L. Wavelet neural network prediction method of stock price trend based on rough set attribute reduction. Appl Soft Comput J. 2018;62:923–32. https://doi.org/10.1016/j.asoc.2017.09.029.

[19] Weng B, Lu L, Wang X, Megahed FM, Martinez W. Predicting short-term stock prices using ensemble methods and online data sources. Expert Syst Appl. 2018;112:258–73. https://doi.org/10.1016/j.eswa.2018.06.016.

[20] Jeon S, Hong B, Chang V. Pattern graph tracking-based stock price prediction using big data. Future Gener Comput Syst. 2018;80:171–87. https://doi.org/10.1016/j.future.2017.02.010.

[21] Pimenta A, Nametala CAL, Guimarães FG, Carrano EG. An automated investing method for stock market based on multiobjective genetic programming. Comput Econ. 2018;52(1):125–44. https://doi.org/10.1007/s10614-017-9665-9.

[22] Billah, M.M., Sultana, A., Bhuiyan, F. *et al.* Stock price prediction: comparison of different moving average techniques using deep learning model. *Neural Comput & Applic* **36**, 5861–5871 (2024). https://doi.org/10.1007/s00521-023-09369-0

[23] Friday, I.K., Pati, S.P., Mishra, D. *et al.* CAGTRADE: Predicting Stock Market Price Movement with a CNN-Attention-GRU Model. *Asia-Pac Financ Markets* (2024). https://doi.org/10.1007/s10690-024-09463-w

[24] Moodi, F. , Jahangard Rafsanjani, A. , Zarifzadeh, S. and Zare Chahooki, M. A. (2024). Advanced Stock Price Forecasting Using a 1D-CNN-GRU-LSTM Model. Journal of AI and Data Mining, 12(3), 393-408. doi: 10.22044/jadm.2024.14831.2581

[25] Li, T., Liu, Z., Shen, Y., Wang, X., Chen, H., & Huang, S. (2024). MASTER: Market-Guided Stock Transformer for Stock Price Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(1), 162-170. https://doi.org/10.1609/aaai.v38i1.27767

[26] Zicheng Tao, Wei Wu, Jianxin Wang, Series decomposition Transformer with period-correlation for stock market index prediction,Expert Systems with Applications, Volume 237, Part B,2024, https://doi.org/10.1016/j.eswa.2023.121424.

[27] Ma, D., & Yuan, D. (2024). Enhanced stock price forecasting through a regularized ensemble framework with graph convolutional networks. *Expert Systems with Applications*, *250*, 123948.

[28] Liao, S., Xie, L., Du, Y., Chen, S., Wan, H., & Xu, H. (2024). Stock trend prediction based on dynamic hypergraph spatio-temporal network. *Applied Soft Computing*, *154*, 111329.