# Improving Upon Multivariate Time Series Forecasting with Graph Neural Networks on Cryptocurrency Datasets

Shriyan Shekhar
Computer Science Department
Hong Kong University of Science and Technology
Hong Kong
sshekharaa@connect.ust.hk

Sung Kit Ko
T&M Department
Hong Kong University of Science and Technology
Hong Kong
skkoab@connect.ust.hk

Siu Hang Tang
Computer Science and Department
Hong Kong University of Science and Technology
Hong Kong
shtangac@connect.ust.hk

***Abstract*—The explosive growth of cryptocurrency markets presents unique challenges for accurate price prediction due to their inherent volatility and complex dependencies. This paper explores the adaptation of Graph Neural Networks (GNNs) for multivariate time series forecasting in the cryptocurrency domain. We propose an enhanced framework, the Multivariate Time-Series Graph Neural Network (MTGNN), which integrates multiple feature processing capabilities and incorporates an attention mechanism to improve prediction accuracy. By leveraging a diverse set of features derived from historical trading data, our model captures both spatial and temporal dependencies effectively. The MTGNN is evaluated against baseline models, including DCRNN, A3TGCN, and T-GCN. Experimental results demonstrate that the MTGNN significantly outperforms traditional models. These findings highlight the potential of GNN-based approaches in enhancing the forecasting capabilities for volatile cryptocurrency markets, paving the way for more informed decision-making in trading.**

*Keywords— GNNs, Spatial and Temporal, Machine Learning*

## I. INTRODUCTION

Trading in cryptocurrency has surged in recent years, driven by its potential for high returns and transformative impact on financial systems. The unique characteristics of cryptocurrencies, including their decentralized nature and innovative technology, present both opportunities and challenges for investors and researchers alike. Understanding and predicting price movements in this dynamic market is crucial for informed decision-making and risk management.

The paper “Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks” [1] presents a novel approach to modeling complex dependencies in time series data. This research aims to adapt the proposed Graph Neural Network (GNN) methodology for forecasting price and trading trends in the cryptocurrency market, a domain characterized by high volatility and intricate spatial and temporal relationships.

Cryptocurrency markets exhibit rapid fluctuations, driven by both short-term events and long-term trends, making accurate prediction challenging. The graph learning layer introduced in MTGNN will allow the model to adaptively learn an appropriate adjacency matrix to capture these hidden dependencies within the data without a need for an initial defined adjacency matrix. Additionally, the proposed mix-hop propagation mechanism enhances model robustness against the inherent volatility of cryptocurrency prices.

To further improve prediction accuracy, integration of an attention mechanism from "Graph Attention Networks" [2] is proposed. This addition allows the model to focus on significant features, increasing resilience against the volatile trends inherent in cryptocurrency datasets. Additionally, effective cryptocurrency prediction necessitates the consideration of multiple features. Therefore, the research incorporates these diverse features into the MTGNN framework, aiming for more accurate price predictions.

## II. METHODOLOGIES

### A. About the Original Work

The main objective of the paper “Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks” [1] is to propose a novel framework that effectively models multivariate time series data by leveraging graph neural networks (GNNs). The paper uses the following procedure:

1. Extracting unidirectional relations among variables through a graph learning module, into which external knowledge like variable attributes can be easily integrated. This is done automatically.

2. Integrate a mix-hop propagation layer and a dilated inception layer to accurately capture both spatial and temporal dependencies.

3. Enable joint learning of the graph structure and predictive modelling in an end-to-end framework.

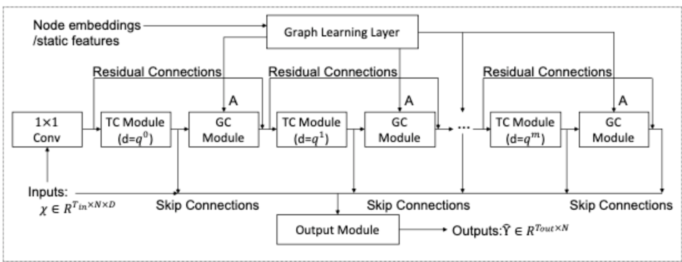


Fig.1. MTGNN Framework

## B. Dataset Overview

The dataset is fetched from the Binance API and consists of historical candlestick data for most notably, BTC/USDT, ETH/USDT, BNB/USDT, and 31 other cryptocurrencies [3]. The timeframe selected for this project spans from October 1, 2019, to October 1, 2024 with hourly and daily timestamps. Each data point provides a 9-feature vector which includes the opening, high, low, and closing prices, as well as volume metrics and trade counts. This structured format allows for comprehensive analysis of price movements and market behavior and further feature engineering.

### B.1. Feature Engineering

The dataset is enhanced with several additional features to provide deeper insights into market dynamics and improve the robustness of the price prediction model. The following features are computed through the existing 9 features from the given dataset and includes:

- Volume Quote Ratio: This metric assesses the relationship between trading volume and quote asset volume, highlighting liquidity conditions.
- Buy-Sell Volume Ratio: This ratio indicates the proportion of buy versus sell volume.
- Buy-Sell Quote Ratio: This ratio focuses on the quote asset volume associated with buy and sell transactions.
- Log Returns: This feature captures the logarithmic returns of prices, providing a normalized view of price changes over time.
- Stochastic Oscillator: This momentum indicator helps identify overbought or oversold conditions in the market.
- Volume-Weighted Average Price (VWAP): VWAP offers a measure of the average price at which an asset has been traded throughout the day, weighted by volume.
- Parkinson's Volatility: This measure estimates volatility based on the high and low prices within a specific period.
- Garman-Klass Volatility: This volatility estimate uses opening, high, low, and closing prices to capture market volatility more accurately.
- Average Size of Trades: This feature reflects the typical size of trades.
- Average Trade Quote Size: This metric assesses the average size of trades in quote asset terms.

These features encompass liquidity metrics, price movement indicators, momentum indicators, volatility estimates, and trade behavior insights. All of which are critical factors that can influence the future price of a cryptocurrency.

## C. About Our MTGNN model – Our Contributions

This research makes several key contributions to the application of Graph Neural Networks (GNNs) for cryptocurrency price prediction in terms of improving the feature space of MTGNN, model architecture, comparison and ablation studies. A significant enhancement to the original MTGNN framework was achieved by integrating these approaches.

By integrating multiple feature processing capabilities, the model is capable of analyzing various features simultaneously rather than being limited to a single input, as implemented in the original research. Additionally, techniques for noise removal and moving average filtering were applied to improve data quality, ensuring that the model operates on cleaner, more relevant information. Hence, standardization and batching processes were refined to accommodate non-stationary characteristics of cryptocurrency market data and resolve possible data leakage in the original implementation.

In terms of model architecture, an attention layer was also added to enhance the model's ability to focus on significant features, potentially improving prediction accuracy. The attention layer is added in between the temporal and graph convolution modules with the following rationale. It elevates prediction accuracy with selective attention on most relevant information of input sequences. Moreover, more global information from the time series would be preserved to address long-term trends within the always volatile cryptocurrency market.

Two versions (oldMTGNN and newMTGNN) of the model and multiple graph-based comparison models were developed, enabling a comprehensive evaluation of different methods. The oldMTGNN is the original implementation developed by the authors of the paper, while the newMTGNN is a PyTorch Geometric implementation. oldMTGNN and newMTGNN differs in the variations of kernel sizes used in dilated inception, with oldMTGNN concatenating convolution results of $1 \times 2$, $1 \times 3$, $1 \times 6$, and $1 \times 7$ kernels, while newMTGNN concatenates results of two $1 \times 1$ kernels. To establish a comparative framework, baseline models from various literature were implemented, providing a basis for assessing the effectiveness of the proposed enhancements. These contributions collectively aim to enhance the accuracy of price predictions in the volatile cryptocurrency market. With these changes, the refined MTGNN architecture is as follows:

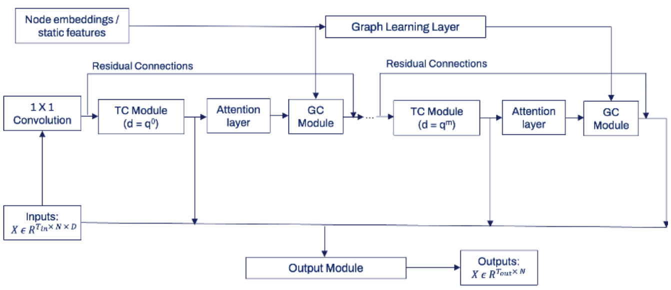


Fig. 2: Refined MTGNN Architecture

### D. Baseline Models

This research aims to compare against three baseline models - DCRNN, A3TGCN, and T-GCN. These models were selected for their ability to accept graph, time-series datasets similar to MTGNN for training and prediction.

The Diffusion Convolutional Recurrent Neural Network (DCRNN) is a model designed for forecasting by effectively capturing spatial and temporal dependencies in time series data. It represents entities as nodes in a directed graph, where the edges are weighted based on relevant metrics that define proximity and interactions. This structure allows for a comprehensive analysis of dynamic processes. The model introduces a diffusion convolution operation that captures spatial dependencies. It employs a sequence-to-sequence architecture to effectively model the temporal aspects of the data, enabling predictions based on historical trends. Additionally, the integration of scheduled sampling enhances the training process by transitioning from ground truth data to the model's own predictions [4].

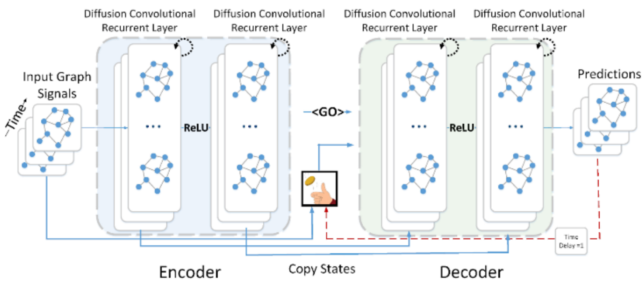


Fig. 3: DCRNN Framework

The Temporal Graph Convolutional Network (T-GCN) is a framework that integrates Graph Convolutional Networks (GCNs) with Gated Recurrent Units (GRUs) to effectively capture spatiotemporal dependencies in dynamic data. By utilizing graph convolution, T-GCN processes the relationships between nodes within a structured graph, enabling it to learn from the underlying topology. Simultaneously, the GRU component addresses temporal dynamics, allowing the model to track changes over time [5]. The model takes in a feature matrix as input, and for each column within the feature matrix, it will pass

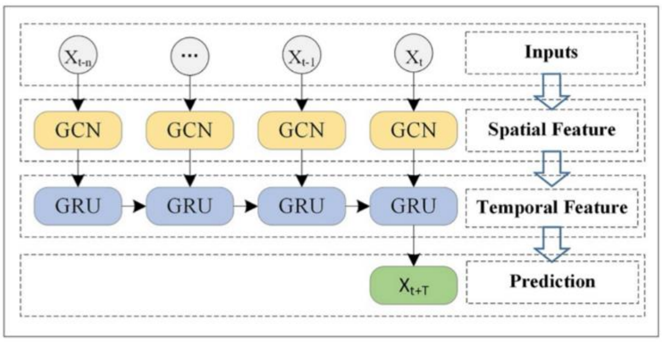


Fig. 4: T-GCN Framework

The Attention-based Temporal Graph Convolutional Network (A3T-GCN) builds upon the T-GCN framework by incorporating an attention mechanism to enhance its predictive capabilities. This model re-weights the significance of historical data points, allowing it to focus on the most relevant information during forecasting tasks. The A3T-GCN captures trends by calculating attention scores for each hidden state derived from the spatiotemporal modeling [6].

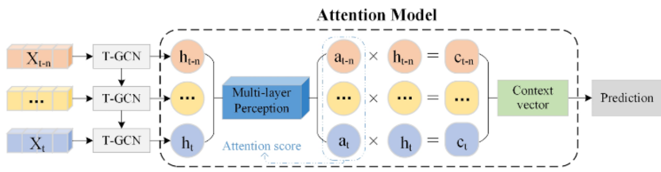


Fig. 5: A3T-GCN Framework

## III. Results

### A. Experimental Settings

All MTGNN variations and baseline models were evaluated under identical experimental settings to ensure a fair comparison of performance. All models were trained on the same cryptocurrency dataset. Two distinct timeframes were utilized for the training sets: one-hour intervals, yielding 27,019 samples per currency, and one-day intervals, resulting in 1,108 samples per currency. A consistent train-validation-test split of 60-20-20 was applied across all experiments, with the training set encompassing the earliest time range of the dataset. Hyperparameters were standardized across models for comparison purposes, which includes a fixed learning rate set to 0.001. The models were trained for 16 epochs using the AdamW optimizer with weight decay set to 0.0005, and a batch size of 16 was employed during training. A single step forecasting horizon of 3 hours/ days was made for predicting close prices in all models. The models were trained on Kaggle using NVIDIA TESLA P100 GPU for training and inference. The number of hidden layers for each of the baseline models will be 2.

A key difference between the baseline models and the MTGNN was the initialization of a graph and adjacency matrix based on the correlation of currencies in the training set time range. The graph learning layer allows the MTGNN to adaptively construct an appropriate adjacency matrix to capture the hidden relationships within the graph. However, this module is non-existent for the baseline models. Therefore, to initialize an adjacency matrix, the correlation between the closing price in the training set of each currency to one another was calculated. Currencies sharing a correlation exceeding 0.7 have an edge initialized between them, otherwise no edge will be constructed. The correlation matrix calculated can be viewed in Figure 6. After constructing the initial edges between cryptocurrencies based on the correlation, an adjacency matrix can be generated. This adjacency matrix will then be used for the baseline models and will act as a substitute for the graph learning layer in the MTGNN framework.

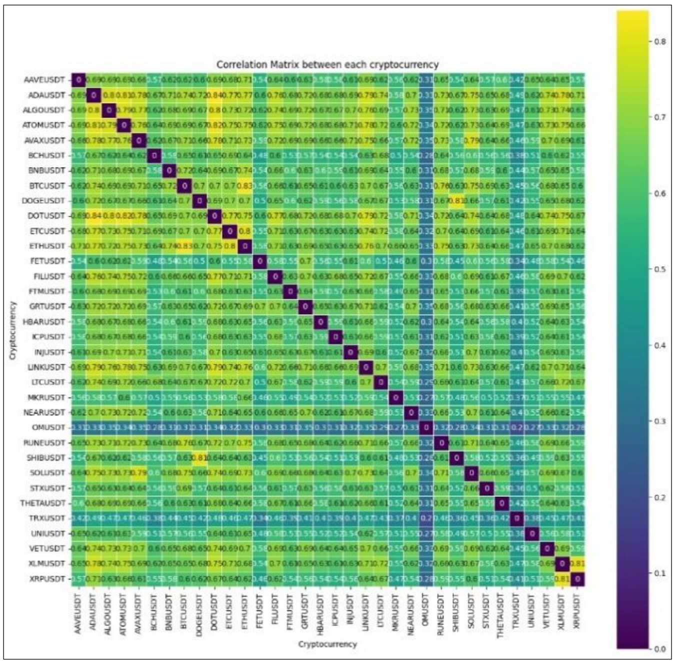

Fig. 6: Correlation Matrix between Cryptocurrencies

## B. *Evaluation Metrics*

To evaluate the performance of the models, the following metrics are used:

- Loss: The total squared differences of prediction errors.
- Root Mean Square Error (RMSE): The average magnitude of the prediction errors.
- Relative Absolute Error (RAE): The accuracy of the model's predictions relative to the absolute error of a naïve Linear model.
- Correlation (CORR): The strength and direction of the linear relationship between the predicted and actual values

## C. *Baseline Model Results*

TABLE I.
BASELINE RESULTS OVER 1 DAY TIMEFRAME

| Models | Test Loss | Test RMSE | Test RAE | Test CORR |
|---|---|---|---|---|
| T-GCN | 4021.89 | 0.7383 | 0.7688 | 0.6758 |
| A3T-GCN | **2543.802** | **0.5872** | **0.6126** | **0.7366** |
| DCRNN | 5374.39 | 0.8535 | 0.9126 | 0.3620 |

As the results from Table I demonstrate, the 1-day timeframe dataset has acceptable performance for T-GCN and A3T-GCN, while DCRNN struggles to learn from the dataset. T-GCN's architecture allows it to capture spatio-temporal dependencies of the provided cryptocurrency space. It is also noticeable that A3T-GCN performs the best in every metric such as the Loss, RMSE, RAE and CORR. The added attention mechanism in A3T-GCN allows the model to improve in performance. Below, is the empirical prediction of the baseline models on ETHUSDT on a daily time interval:

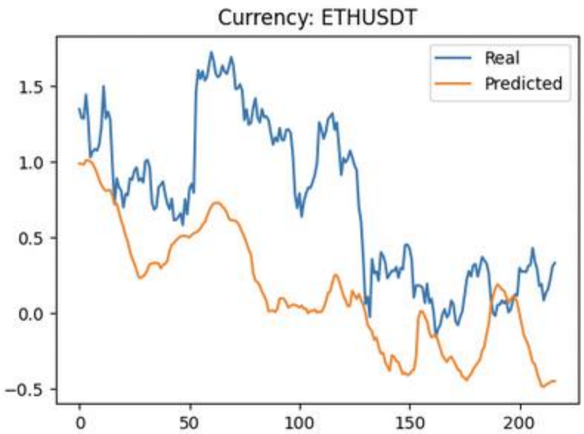

Fig. 7: A3T-GCN Prediction for ETHUSDT-1d

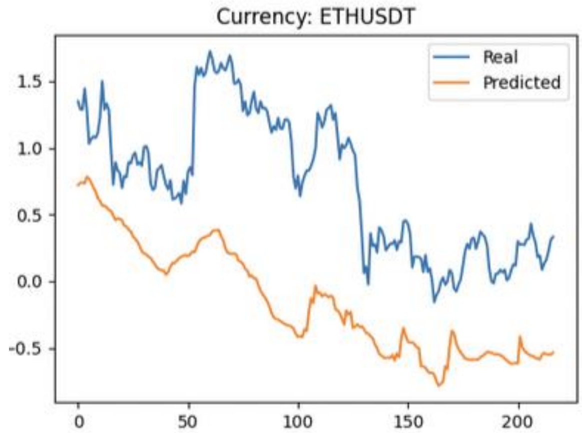

Fig. 8: T-GCN Prediction for ETHUSDT-1d

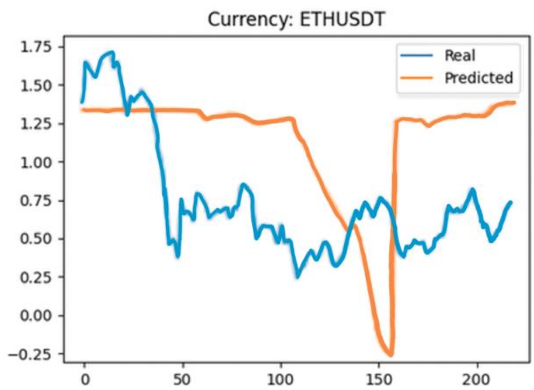

Fig. 9: DCRNN Prediction for ETHUSDT-1d

For currencies with large prices such as ETHUSDT, the baseline models will struggle in predicting the prices for daily timeframe. Although, there are instances where the model is able to capture the overall trend of the currency’s prices, but not the short-term fluctuations.

TABLE II.
BASELINE RESULTS OVER 1 HOUR TIMEFRAME

| Models | Test Loss | Test RMSE | Test RAE | Test CORR |
|---|---|---|---|---|
| T-GCN | **7871.092** | **0.2928** | **0.1772** | 0.9360 |
| A3T-GCN | 8497.078 | 0.3042 | 0.2328 | **0.9417** |
| DCRNN | 51904.086 | 0.7519 | 0.8107 | 0.5655 |

Evidently, as seen in Table II, the performance every baseline model improved significantly when using hourly timeframe. Shorter timeframe means that the models can learn the short-term dependencies of each currency and capture trends more accurately, and a smaller timeframe means a larger dataset (number of samples increased from 1,108 to 27,019) which will improve the model's prediction capabilities.

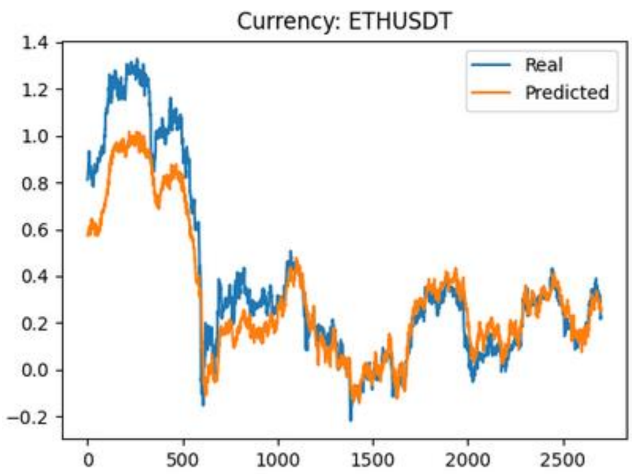


Fig. 10: A3T-GCN Prediction for ETHUSDT-1h

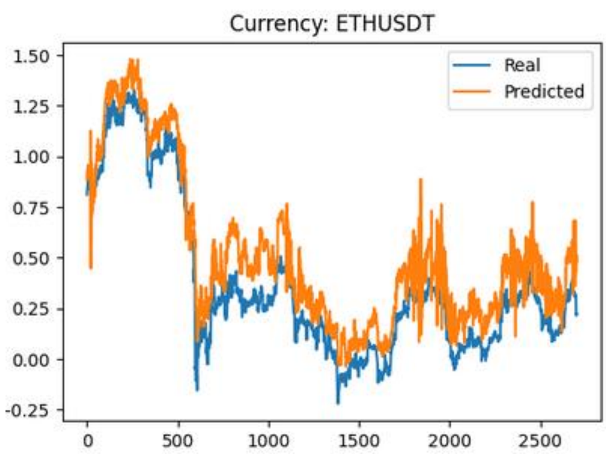


Fig. 11: T-GCN Prediction for ETHUSDT-1h

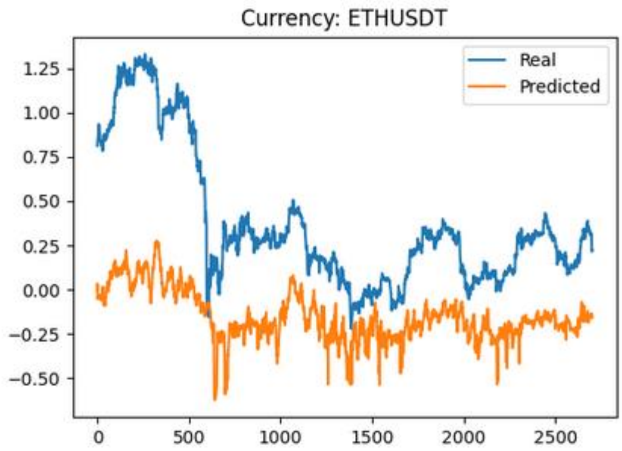


Fig. 12: DCRNN Prediction for ETHUSDT-1h

As the charts display, T-GCN and A3T-GCN were able to capture the long-term trends and short-term fluctuations in normalized closing prices for ETHUSDT quite well when using hourly time interval, where the latter has the smaller margin of error. However, DCRNN continues to underperform, where now it is able to capture the fluctuations in prices, however, it still struggles to accurately predict the price range of the currencies.

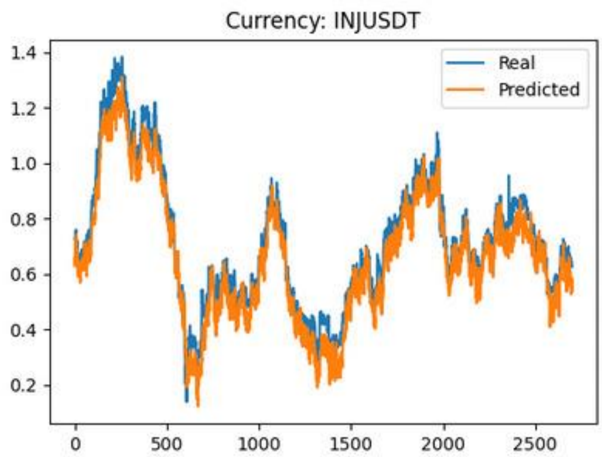


Fig. 13: A3T-GCN Best Prediction

Figure 13 illustrates the most accurate prediction made for a cryptocurrency by the A3T-GCN model, specifically for INJUSTD cryptocurrency. The figure clearly shows a close alignment between the actual and predicted values, reinforcing the assertion that A3T-GCN is the best baseline model out of the following.

### D. *MTGNN Results*

TABLE I.
COMPARING VARIOUS IMPLEMENTED MTGNN MODELS

| Run details | Test Loss | Test RMSE | Test RAE | Test CORR |
|---|---|---|---|---|
| Old - All Features 1h | **686.4324** | **0.0024** | **0.0002** | 0.9950 |
| New - All Features with Attention 1h | 1281.4488 | 0.0846 | 0.0222 | 0.9943 |
| New - One Feature with Attention 1h | 1300.2911 | 0.0853 | 0.0223 | 0.9939 |
| New - All Features with Moving Average 1h | 1335.9319 | 0.0864 | 0.0230 | 0.9879 |
| New – All Features 1h | 1343.9429 | 0.0867 | 0.0232 | 0.9920 |
| Old - One Feature with Moving Average 1h | 1344.6059 | 0.0868 | 0.0230 | 0.9869 |
| New - One Feature with Moving Average 1h | 1345.3467 | 0.0868 | 0.0230 | 0.9862 |
| New - One Feature with Moving Average and Attention 1h | 1346.5454 | 0.0869 | 0.0232 | 0.9873 |
| Old - One Feature 1h | 1349.2517 | 0.0869 | 0.0232 | **0.9980** |
| New - One Feature 1h | 1349.6737 | 0.0870 | 0.0224 | 0.9947 |
| Old - One Feature 1d | 31486.99 | 2.0517 | 0.5485 | 0.7394 |
| New – One Feature 1d | 33005.841 | 2.1006 | 0.5606 | 0.6675 |

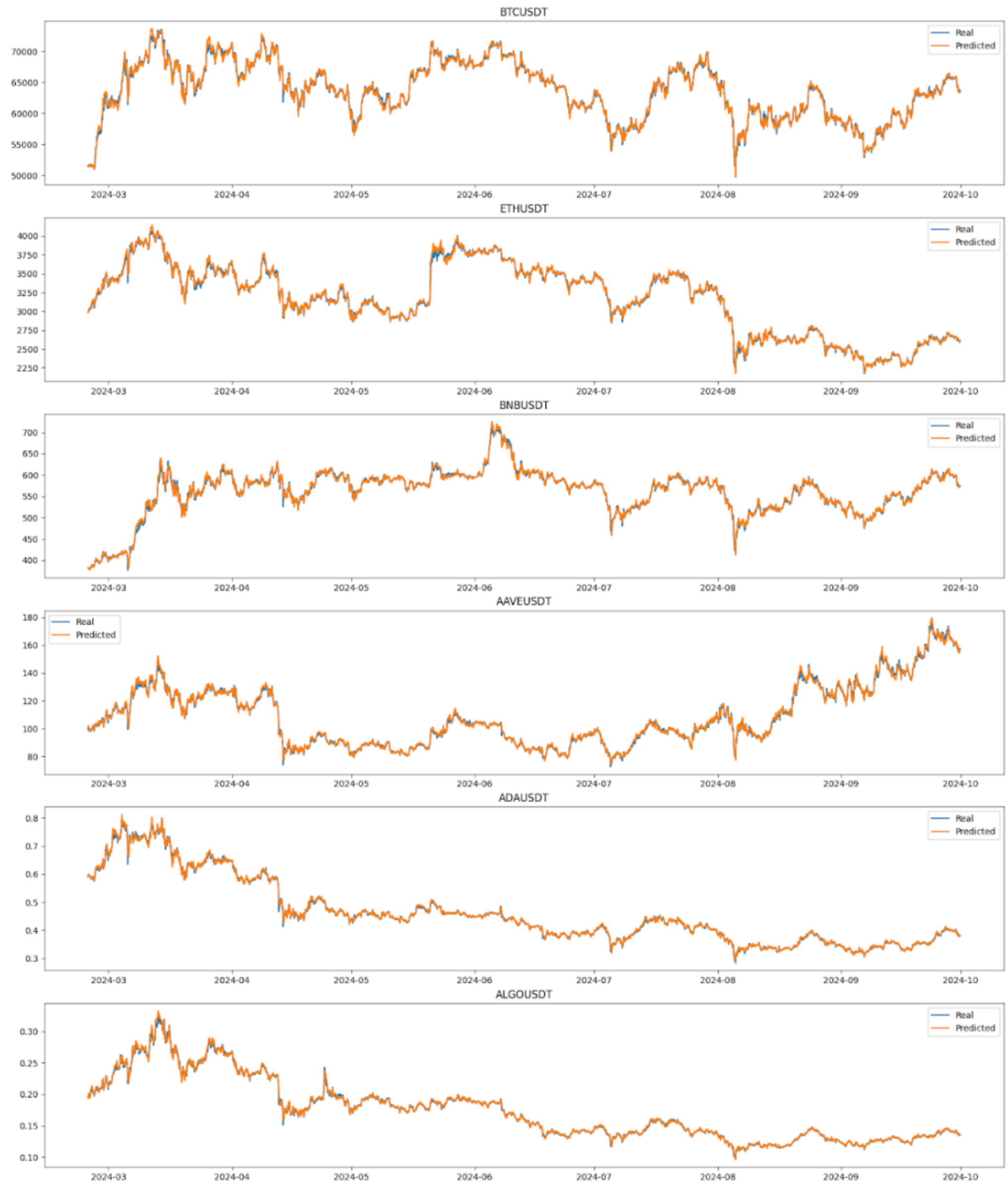


Fig 14. Comparison of Predicted vs. Actual Values for Six Cryptocurrencies Using Best MTGNN – Old All Features 1h

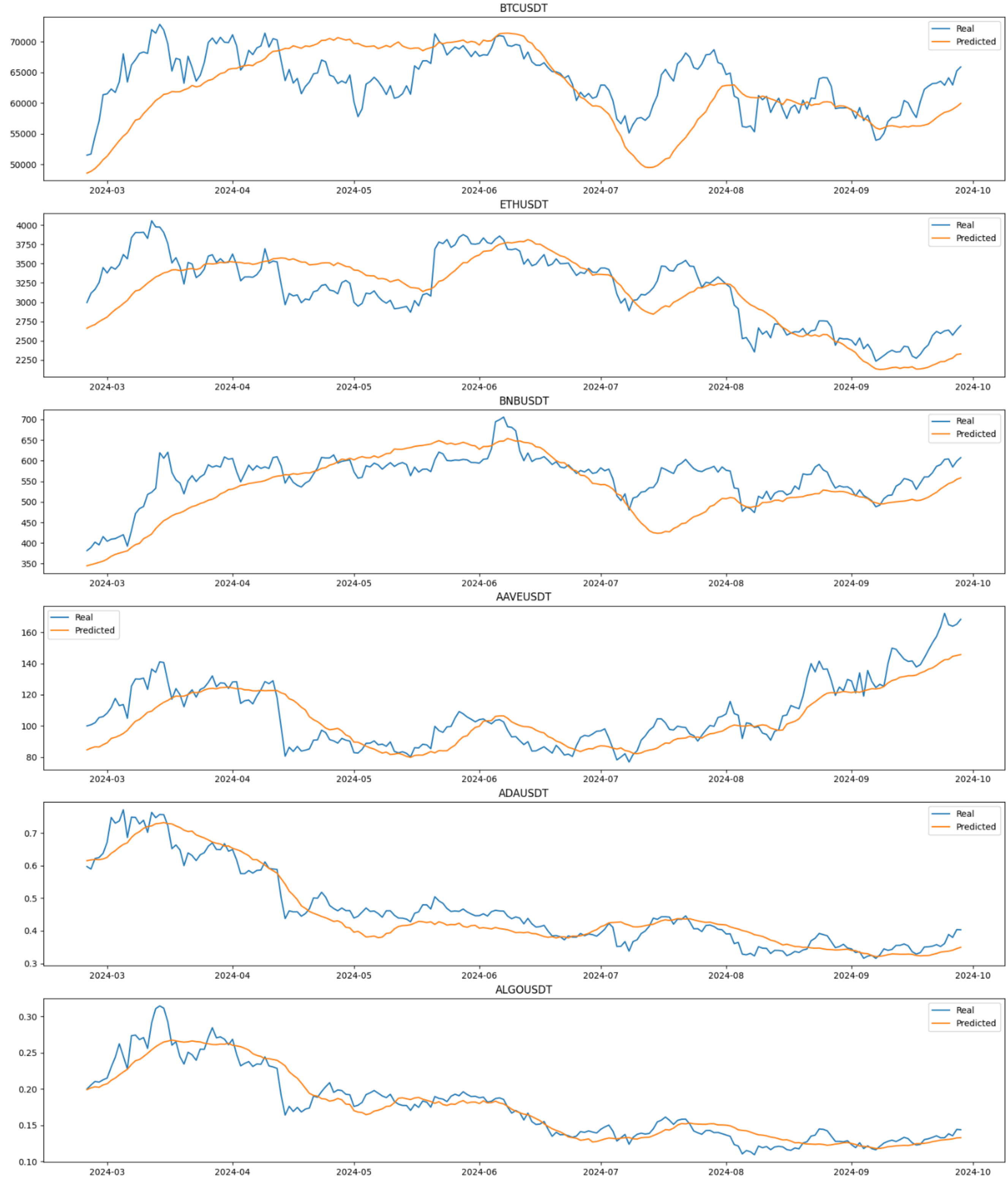


Fig 15. Comparison of Predicted vs. Actual Values for Six Cryptocurrencies Using Best MTGNN – Old One Feature 1d

## E. Insights

In this subsection, the comparison between different settings of the MTGNN with other graph-based time-series prediction models was performed. Tables I and II show the experimental results of three proposed baseline models, differ in the data timeframe used in training. Table III shows the experimental results of different implementations of the MTGNN model. Overall, the MTGNN model successfully outperformed across all metrics of the baseline models for hourly timeframe. Moreover, it is noted that hourly data outperformed daily data significantly as noticed by differences in Figure 14 and 15, which can be explained by insufficient training examples of daily cryptocurrency data, lacking data depth to keep track of spatial and temporal dependencies. In particular, the original MTGNN model trained with hourly data, consisting of multiple added features, performed the best out of all runs. The main reason behind the successful improvement of MTGNN is the addition of multiple price and volume related features. Comprehensive spatial dependencies can be captured through temporal convolutions and improve prediction accuracy. Additionally, the original implementation outperforms the PyTorch Geometric implementation. The main reason behind the difference in variations of aforementioned kernel sizes is in the dilated inception layers. More variations provided in the original implementation means more comprehensive dependencies being captured in temporal convolution modules. Therefore, this implementation provides the best result out of all runs. The outcomes of this particular run are illustrated in Figure 14, which shows the performance across the first six cryptocurrencies. In this run, the actual and predicted prices are noticeably highly correlated, which validates the findings presented Table III for this run.

### E.1 Insights – Compare

Comparisons between the results obtained from Table I, Table II and Table III show the improvements the MTGNN models made on predicting the dataset for hourly timeframe compared to the baseline models. Test RMSE for MTGNN for hourly data is always < 0.1 while the baseline models have errors > 0.3. It is apparent that the additional modules present in the MTGNN, such as the graph learning layer allowing the model to adaptively learn an appropriate adjacency matrix is more advantageous than initializing a fixed adjacency matrix for the baseline models, which may not properly model the relationships between nodes/cryptocurrencies appropriately. This also might be due to the model complexity of MTGNN, the baseline models have comparatively few parameters; Given that the number of hidden layers for the baseline models are 2, T-GCN has 5,809, A3T-GCN has 5,826, and DCRNN has 4,817. Which is significantly lower than the new MTGNN's 393,825 parameters. More trainable parameters means that the model is able to capture more inherent patterns but could risk overfitting and extending training time significantly. This could explain the difference in accuracy between the MTGNN and the baseline models.

However, in the case of the daily timeframe dataset, the MTGNN models (only using one feature) perform slightly worse than T-GCN and A3T-GCN as seen from the results shown in Table II. The MTGNN model was able to learn the long-term trends of the currencies from Figure 15, but the baseline models (excluding DCRNN) were able to better model the short-term fluctuations when analyzing stablecoins. For currencies with higher volatility and prices such as BTCUSDT and ETHUSDT, the baseline models struggle to predict the short-term and long-term dependencies of the price, while the MTGNN performs better in modelling the long-term trends of each of those coins, meaning MTGNN was able to learn the long-term trends in comparison to the baseline models. This shows that MTGNN is more robust to smaller dataset sizes and limited features since the results displayed use the old implementation which only accepts one feature per node for training. This could be due to either the additional modules detailed previously, or the larger parameter size.

## F. Ablation Studies

An ablation study is conducted to validate the effectiveness of key components that contribute to the improvement of the MTGNN model. The names of the MTGNN with different components are as follows:

- Only Multiple Features: MTGNN with building additional features only, including different price and volatility indicators.
- Only Noise Removal: MTGNN with the usage of moving average filters for noise removal only, removing noise in data for prediction directly.
- Only Attention Layer: MTGNN with the addition of attention layer between temporal and graph convolution modules only.
- Attention Layer and Noise Removal: MTGNN with noise removal in preprocessing, and attention layer added to model, containing close prices of each cryptocurrency only.
- Multiple Features and Attention: MTGNN with multiple features and attention layer added between the aforementioned layers.
- Multiple Features and Noise Removal: MTGNN with multiple features and noise removal in data preprocessing.
- Multiple Features, Noise Removal and Attention: MTGNN with all additional changes proposed in this research.

The study is conducted by replicating the settings in the MTGNN experiment, reporting the same evaluation metrics as stated in Table II. The introduction of multiple features significantly improves the result as it increases the exploration depth and comprehensiveness. The introduction of noise removal in preprocessing is evident as well. Moving average filters can effectively remove random noise in data to identify encoded in the time domain while maintaining short training time with ease of implementation. The attention layer is advantageous to error reduction as well by introducing weights for selective features for graph convolutions, providing an additional filter after temporal convolution. By integrating all possible variations of proposed changes, it is noted that multiple features tend to be the necessary change to improve model performance. One out of noise removal and attention layers can be introduced to improve prediction accuracy, but not both to avoid over-smoothing and produce higher test errors.

## IV. Limitations and Future scope

### A. Limitations

The charts demonstrate that the actual and predicted are quite similar, however, there are some instances where the predicted and actual have a 5% difference. This is evident in Bitcoin prediction. Although the difference percentage appears to be quite low, Bitcoin's value is approximately 55000 USD therefore, an inaccurate prediction of 5% can result in thousands of dollars deviation from actual.

Additionally, the dataset is quite large, encompassing numerous parameters and features. In fact, the new MTGNN model utilizes 393,825 parameters. The computational runtime for a complete run of MTGNN is around 12 hours when processing multiple features. This extensive processing time highlights the complexity involved and leads to all runs for 1d timeframe to be incomplete.

Lastly, cryptocurrency prices are heavily influenced by market sentiment, news events, and social media trends. These factors can be unpredictable and difficult to quantify, making it challenging for our model to account for sudden price shifts. This research lacks market sentiment analysis as the dataset was not available for all the cryptocurrencies our team aimed to assess and use for prediction.

### B. Future Scope

In the future, we plan to enhance our predictive models by employing more sophisticated ensemble methods. Specifically, we aim to develop a meta-model that leverages the prediction results from various models operating over different timeframes. This approach will allow us to capture a wider range of market behaviors and improve overall accuracy. By integrating predictions from multiple models, we can mitigate individual model biases towards particular time ranges and improve handling of complexities inherent in cryptocurrency price movements. This strategy, in conjunction with our current use of moving averages to capture long-term dependencies, will help create a more robust and adaptable forecasting framework

Additionally, with better computational power, new cryptocurrencies and expanded historical range can be added to the dataset to improve the model's predictive capabilities. This expansion will enhance the model's ability to analyze relationships and trends across a broader spectrum of digital assets. Eventually, this will lead to better generalization and insights. The features for the dataset can also be expanded with market sentiment and qualitative features through sources such as social media and news articles, in order to incorporate consumer's expectations into our model to expand the feature space, allowing the model to consider qualitative variables that potentially affect currency prices. Furthermore, sophisticated back-testing can be performed to validate the predictive capabilities of MTGNN on historical data, in order to analyze the model's capabilities for real-world applications.

## V. Conclusion

In this study, we have introduced the Multivariate Time-Series Graph Neural Network (MTGNN) as a novel approach for predicting cryptocurrency prices, addressing the unique challenges posed by the volatile nature of these markets. By integrating multiple feature processing capabilities and an attention mechanism, our model effectively captures both spatial and temporal dependencies inherent in cryptocurrency data.

Our experimental results demonstrated that the MTGNN significantly outperforms established baseline models, including DCRNN, A3TGCN, and T-GCN, particularly when utilizing shorter time frames such as 1h. The incorporation of diverse features, alongside preprocessing techniques such as noise removal, has proven critical in enhancing prediction accuracy. Additionally, the attention layer has facilitated a more focused analysis of significant features, further improving the model's performance.

Despite these advancements, our research acknowledges certain limitations, including the influence of market sentiment and the extensive computational resources required for model training. These challenges highlight the need for ongoing refinement and adaptation of our methodologies.

Looking ahead, we envision enhancing our predictive framework by exploring ensemble methods that combine predictions from various models. Such developments will not only improve accuracy but also provide a more comprehensive understanding of the cryptocurrency landscape. Expanding our dataset to include a wider array of cryptocurrencies will further enrich our analysis, enabling us to capture complex interdependencies between cryptocurrencies and enhance our forecasts. Ultimately, this research contributes to the growing body of knowledge in cryptocurrency prediction, supporting more informed decision-making in trading and investment strategies.

## References


[1] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, Xiaojun Chang, and Chengqi Zhang. 2020. Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks. In 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '20), August 23–27, 2020, Virtual Event, USA. ACM, New York, NY, USA, 11 pages. https://arxiv.org/pdf/2005.11650

[2] Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph attention networks. arXiv preprint arXiv:1710.10903. 2017 Oct 30

[3] Binance. Rest API. Binance Open https://developers.binance.com/docs/binance-spot-api-docs/rest-api#klines Platform. (n.d.)

[4] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting," *arXiv preprint arXiv:1707.01926*, 2018.

[5] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, "T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction," *arXiv preprint arXiv:1811.05320*, 2019.

[6] J. Zhu, Y. Song, L. Zhao, and H. Li, "A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting," *arXiv preprint arXiv:2006.11583*, 2020.