MATH 5470 Final-Project: G-Research Crypto Forecasting

Hu Bo[1], Wu Hongfan[1], Qiu wenxi[1] and WANG Zetao[2] {bhuan, hwucn, wqiuae, zetao.wang}@connect.ust.hk
[1]: Department of IEDA, HKUST [2]: Department of Civil and Environmental Engineering, HKUST

1. Introduction

Cryptocurrency, as an increasingly hot investment target, has gradually become the focus of research in price trend prediction. Considering the characteristic of cryptocurrency prices as a time-series data, we employed deep learning methods for price forecasting. Specifically, we constructed the training models using the LSTM model and ResNet residual network. To obtain more effective predictions, we first used the price of the currency as the prediction target. Furthermore, combined with comprehensive data analysis, we utilized target data excluding the influence of market signals as our prediction target. And we have also tested different model struture to gain better results. Finally, we analyzed the results of these experiments.

2. Data Analysis

- This dataset contains information on historic trades for several cryptoassets, such as Bitcoin and Ethereum.
- The detailed data structure is shown as follow:

	timestamp	Asset_ID	Count	Open	High	Low	Close	Volume	VWAP	Target
0	1514764860	2	40.0	2376.580000	2399.5000	2357.1400	2374.590000	19.233005	2373.116392	-0.004218
1	1514764860	0	5.0	8.530000	8.5300	8.5300	8.530000	78.380000	8.530000	-0.014399

- **Timestamp**: Indicating minute-by-minute data.
- **Asset_ID**: All 14 diffierent crpytocurriencies.
- **Count**: Total number of trades.
- **VWAP**:Aggregated form of trade data.
- **Target**: Residual log-returns over a 15 minute horizon.

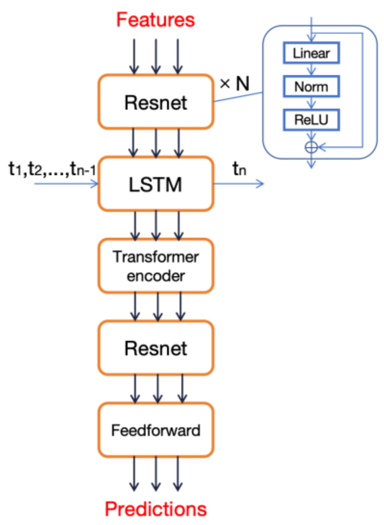
Candlestick chart of Bitcoin price

Analysis

- The price of the cryptocurriencies is a time-series based dataset, the hierarchical time series model may be useful.
- The volatility and correlation structure in the data are likely to be highly non-stationary.
- Due to changes in the overall market environment, fluctuations in currency value may be misleading.

3. Network Structure

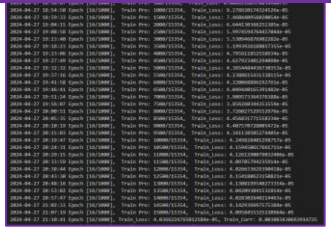
First we used LSTM model in order to guarantee that the features of the time series data can be extracted to make better predictions. We started by splitting the data into segments of 128 units in length, feeding them into the model and outputting the same length of predicted data. However, the results showed that the results of such a training model structure were not satisfactory, so we further thought about the means of improvement by using a training length segmentation of 128 units of data while outputting multidimensional data of one unit in length. The new model showed significant improvement in the validation set.



Model Structure

4. Prediction Results of Different Methods

Without target feature, long output: correlation 0.001
With target feature, long output: correlation 0.002
With target feature, short output: correlation 0.003



The Prediction Result

5. Analysis

Initially, we used a CNN-LSTM-Transformer model, but found that the experimental results were not satisfactory. Then, we replaced the CNN with a multi-layer Resnet model for better feature extraction. During the training process, we continuously updated the learning rate to allow the model to converge to better results. Simultaneously, we adjusted the number of layers in both the Resnet and the Transformer encoder. Regarding the LSTM structure, at first, we input an entire sequence and the LSTM predicted the current time step for each time step in the sequence, but we found the performance unsatisfactory. Ultimately, after referring to papers and code, we modified the model to predict only the next time step following the input sequence. Using this model structure, our correlation could reach 0.003.

6. Conclusion

Through continuously attempting different model architectures, optimizing hyperparameters, and referring to advanced methods, we successfully improved the model's performance. Particularly, by adopting Resnet for feature extraction and adjusting the LSTM's prediction method, the correlation metric significantly improved, reaching 0.003. This reflects the model's excellent performance in feature extraction and sequence prediction, laying the foundation for further model optimization.

7. Contribution

Model construction and training
- HU Bo, Wu Hongfan

Analysis
- HU Bo

Poster written
- Wu Hongfan, Qiu Wenxi, Wang Zetao