Deep Learning-Based Cryptocurrency Price Prediction in Relation to Trading Volume

Ilia Gershenzon | BSc. (Hons.) Computer Science | Lancaster University

Overview

Cryptocurrency:

Total market value is over $942 billion as of October 2022 [1]

Rapid growth attracts new investors

High volatility puts invested money at risk

Prediction of price is beneficial for individual investors and financial researchers [2]

Trading Volume:

Trading Volume is visibly overlooked as a parameter in Cryptocurrency prediction research [3]

Deep Learning (DL):

Cryptocurrencies do demonstrate non-linear patterns in price behaviour

Hence, machine learning tools perform imperfectly [2]

There is a need for a more powerful prediction tool - DL

DL is a well-known solution for complicated time-series problems [3]

Hybrid DL models:

HDL models outperform DL and ML solution for time-series problems

HDL models state of the art solution for cryptocurrency price prediction [4]

HDL models are relevant to the current research [3]

Transformer Model:

Transformer Models got very popular in past year (ChatGPT)

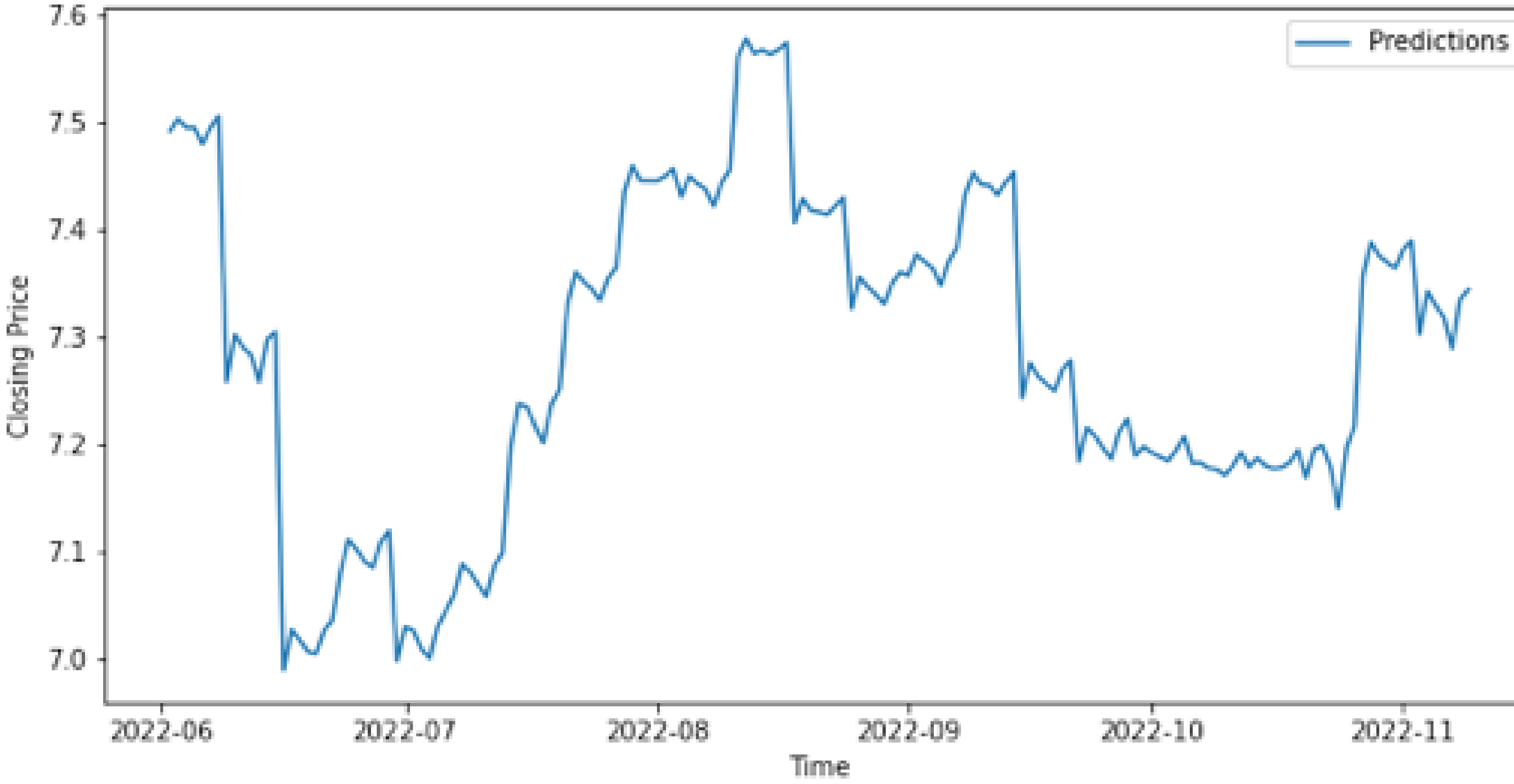
Outperformed old leaders in most NLP tasks [6]

They were never used solely for cryptocurrency price prediction

Maybe there is a hidden prediction power that can compete with current leaders in time-series forecasting

Overview

HDL Methodology



Initial:

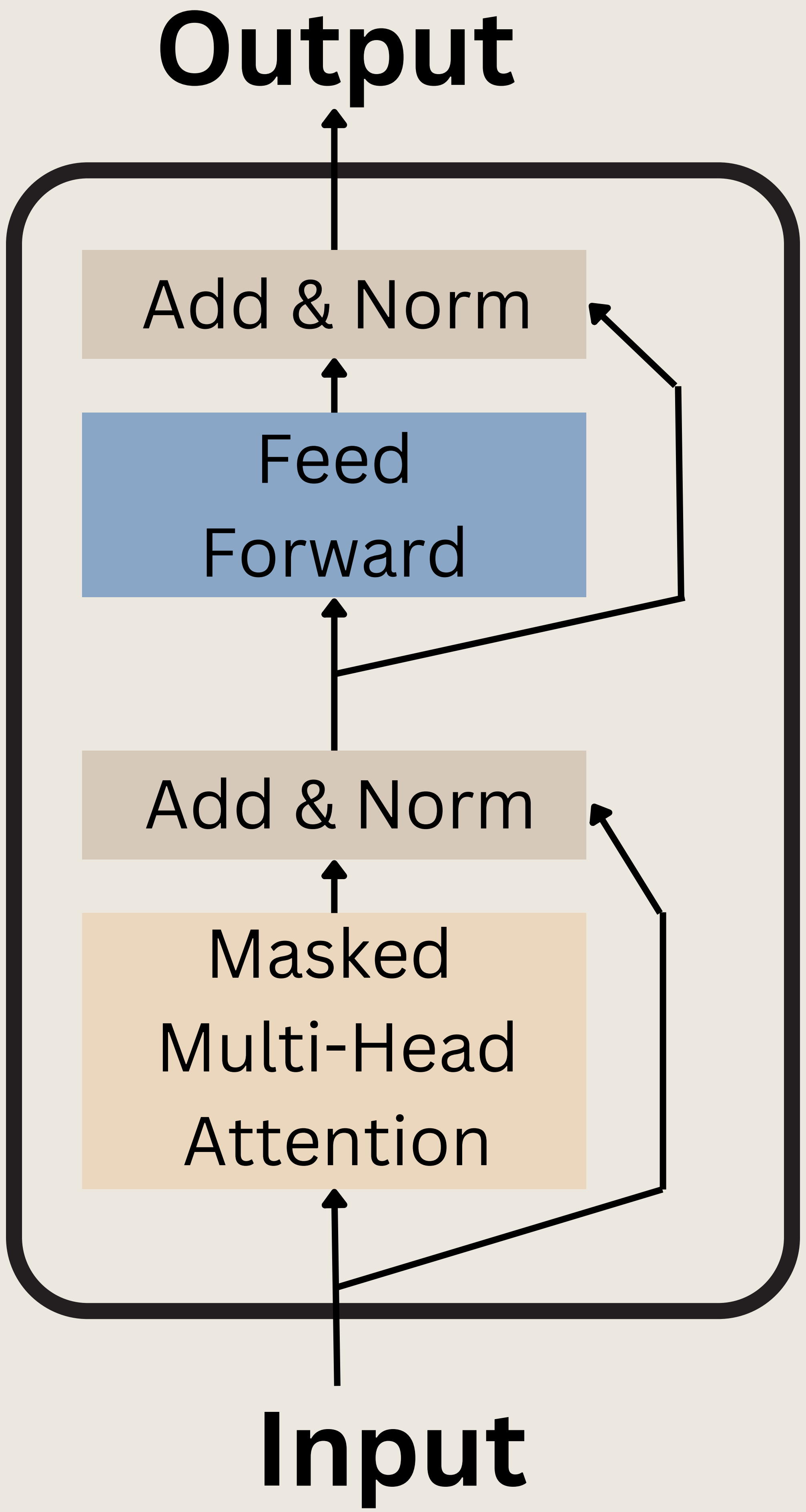
- LSTM-GRU and GRU-LSTM [4]
- Initial design [5]
- Output window -> One week
- Input window experiment -> One week
- Normalisation experiment -> Log scaling

With T Volume:

- Same models
- Different input 7,1 -> 7,2
- Normalisation experiment -> Log scaling

Transformers Methodology

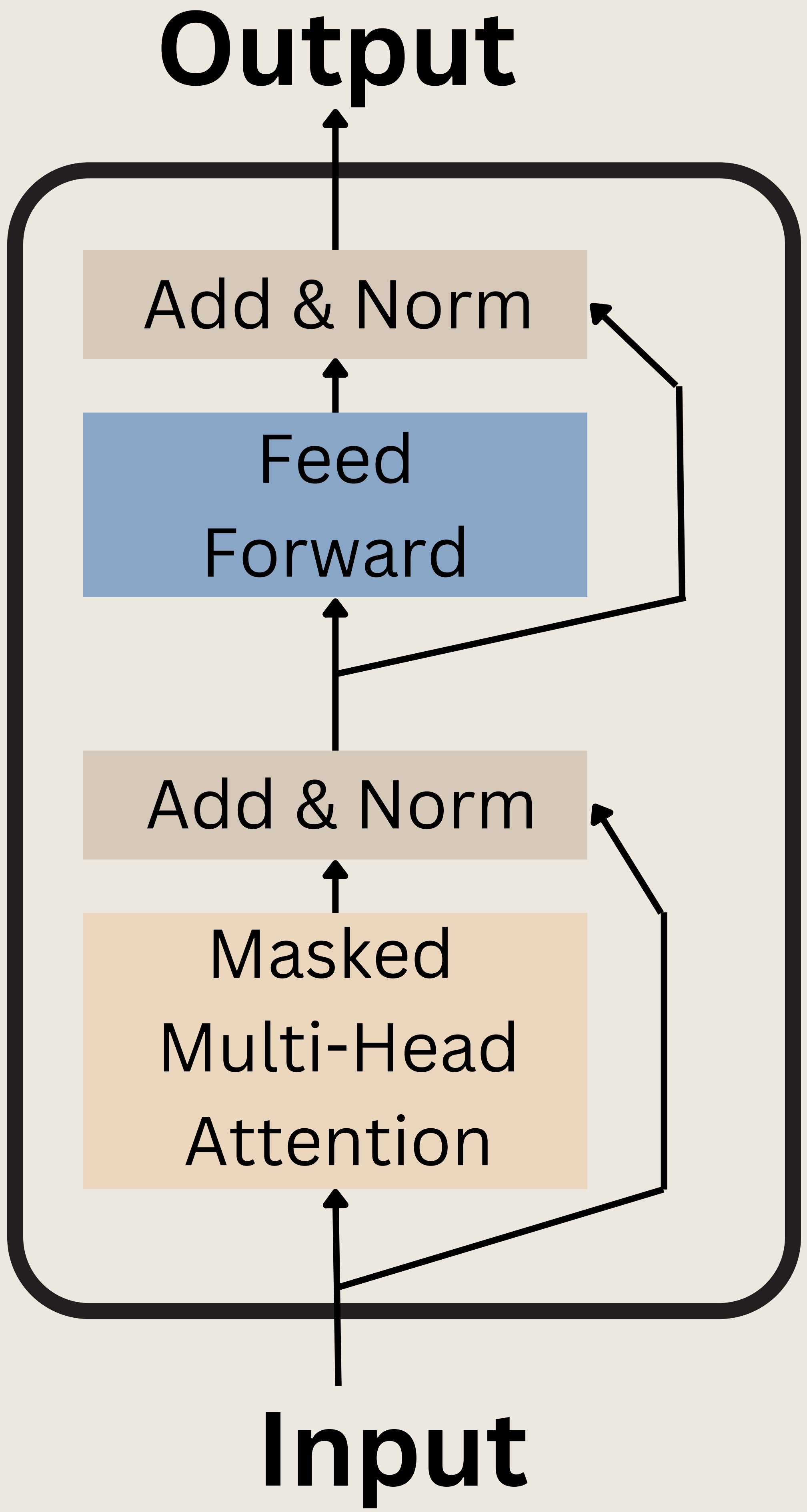
- Decoder only transformer.
- Based on concepts from Attention Is All You Need [6], Language Models are Few-Shot Learners [7].
- Input window experiment -> One week
- Normalisation experiment -> Log scaling



Transformers Methodology

Following elements were developed:

- Attention, Multi-Head Attention, Masked Multi-Head Attention
- Feed Forward block
- Residual blocks
- Dropout
- Custom encoding + decoding



Transformers Methodology
With Volume

- In HDL input for this part is 2d, for transformer input is a mixture of Closing price and Volume.

- Normalisation used: Log scaling.
- One week input, one week output.
- An accuracy drop was expected and happened.

	Log scaling	Volume Log scaling
TM	MSE: **2.6672**	MSE: 134.58
	MAE: **1.3921**	MAE: 8.89
	RMSE: **1.6331**	RMSE: 11.60
	MAPE: **19.06**	MAPE: 122.66

Overview

Results for HDL Models

1. Min-Max with volume -> drop in accuracy 20-32% across different metrics.
2. Log scaling with volume -> a noticeable increase in accuracy.

	Log scaling	Volume + Log scaling	Positive % difference
LSTM-GRU	MSE: 0.0214	MSE: **0.0199**	7.2
	MAE: 0.1058	MAE: **0.1055**	0.3
	RMSE: 0.1463	RMSE: **0.1413**	3.4
	MAPE: 1.47	MAPE: **1.46**	0.7
GRU-LSTM	MSE: 0.0253	MSE: **0.0228**	10.4
	MAE: 0.1175	MAE: **0.1070**	9.4
	RMSE: 0.1589	RMSE: **0.1513**	4.9
	MAPE: 1.63	MAPE: **1.48**	9.6

Results for Transformers Models

1. Volume had negative impact on performance.
2. Generally performance with Log scaling normalisation is promising.
   (In level with HDL models with 2-3 weeks input window)

	Log scaling
TM	MSE: 2.6672
	MAE: 1.3921
	RMSE: 1.6331
	MAPE: 19.06

Conclusion

1. Addition of volume is good for HDL models with Log-scaling normalisation.
2. Addition of volume is harmful in case of Transformer model.
3. Transformer model has room for improvement.

Further Research

Possible ways to improve Transformer:

- Increasing size of the model
- Training on a larger data set
- Fine-tuning to specific task [7]

Thank You

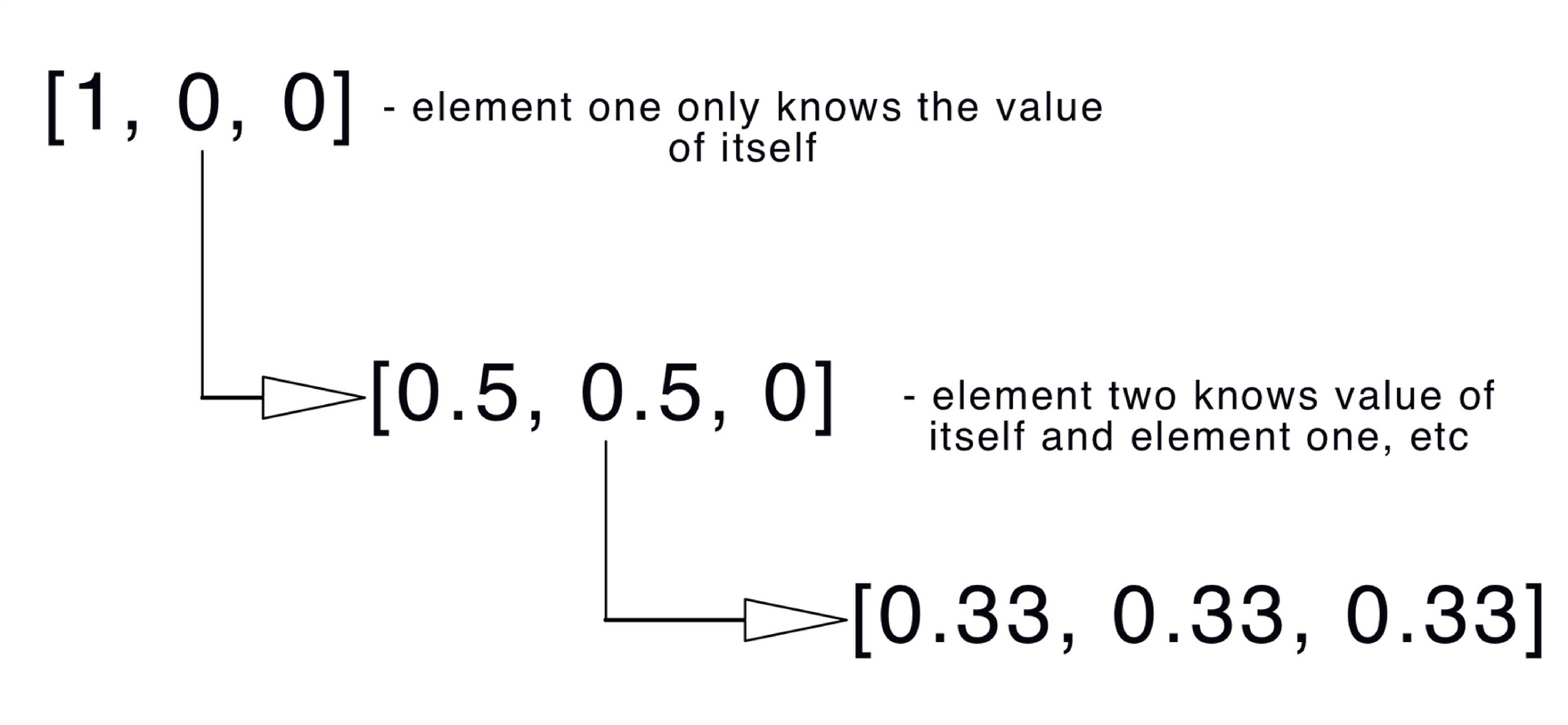
Presented by Ilia Gershenzon

References

[1] CoinMarketCap, “Cryptocurrency Market Capitalizations | CoinMarketCap,” CoinMarketCap, 2019. https://coinmarketcap.com
Accessed on 24th October 2022.
[2] E. Pintelas, I. E. Livieris, S. Stavroyiannis, T. Kotsilieris, and P. Pintelas, “Investigating the Problem of Cryptocurrency Price Prediction: A Deep Learning Approach,” Artificial Intelligence Applications and Innovations, vol. 584, pp. 99–110, May 2020, doi: 10.1007/978-3-030-49186-4_9.
Accessed on 24th October 2022.
[3] A. M. Khedr, I. Arif, P. R. P V, M. El-Bannany, S. M. Alhashmi, and M. Sreedharan, “Cryptocurrency price prediction using traditional statistical and machine-learning techniques: A survey,” Intelligent Systems in Accounting, Finance and Management, vol. 28, no. 1, pp. 3–34, Jan. 2021, doi: 10.1002/isaf.1488.
Accessed on 23th October 2022.
[4] A. Politis, K. Doka, and N. Koziris, “Ether Price Prediction Using Advanced Deep Learning Models,” IEEE Xplore, May 01, 2021. https://ieeexplore.ieee.org/document/9461061/metrics#metrics
Accessed on 22th October 2022.
[5] J. Sen and S. Mehtab, “A Time Series Analysis-Based Stock Price Prediction Using Machine Learning and Deep Learning Models,” International Journal of Business Forecasting and Marketing Intelligence, vol. 6, no. 1, p. 272, 2020, doi: https://doi.org/10.1504/ijbfmi.2020.10038524.
[6] A. Vaswani et al., “Attention Is All You Need,” arXiv preprint, doi: 10.48550/arXiv.1706.03762, Dec. 2017.
Accessed on 10th February 2023.
[7] T. Brown et al., “Language Models are Few-Shot Learners,” arXiv preprint, doi: 10.48550/arXiv.2005.14165, Jul. 2020. Accessed: Feb. 12, 2023. [Online]. Available: https://arxiv.org/pdf/2005.14165.pdf

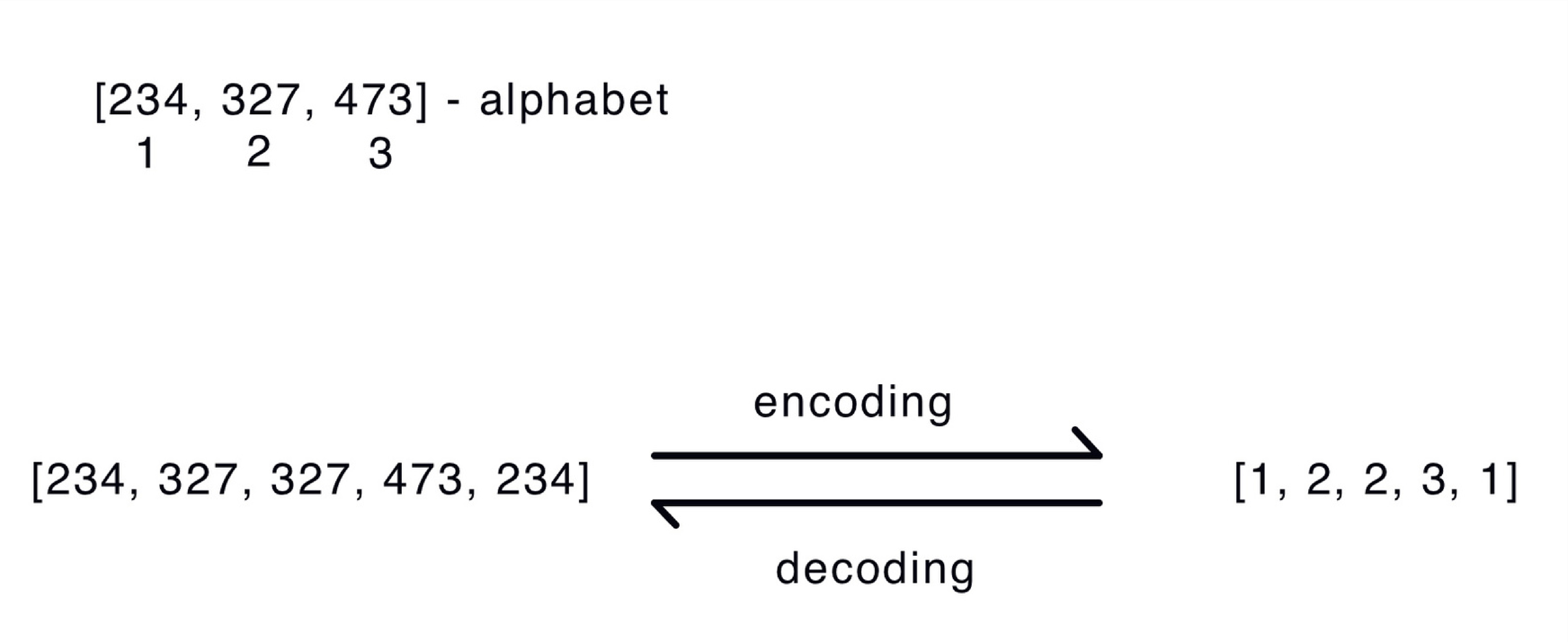
Additional information

Triangular mask example



Additional information

Encoder decoder example



Additional information

Masked Multi-Head Attention