

Insert Title Here

by

PARBHU Varun

Project Supervisor: Professor (Dr) BHURUTH MUDDUN OSK

Project submitted to the Department of Mathematics,

Faculty of Science, University of Mauritius,

as a partial fulfillment of the requirement

for the degree of

M.Sc. Mathematical and Scientific Computing (Part time)

December 2022

Contents

List of Figures	ii
List of Tables	iii
Acknowledgement	iv
Abstract	v
Terms and Definitions	vi
1 Introduction	1
2 Conclusion and Future Works	4

List of Figures

List of Tables

Acknowledgement

First and foremost, I would like to show my deepest gratitude to my supervisor Professor (Dr) M. Bhuruth for his guidance and recommendations throughout the project. Also, I would like to thank my family and friends for their constant support during the coursework.

Lastly, I would like to thank my colleagues for their constant support.

Abstract

Placeholder

Terms and Definitions

α	Learning Rate
η	Reduced Learning Rate
\mathbf{I}	Identity Matrix
\odot	Hadamard Product
NN	Neural Network
LSTM	Long Short-Term Memory
MSE	Mean Squared Error
MAE	Mean Absolute Error
NLP	Natural Language Processing

Chapter 1

Introduction

A cryptocurrency is a decentralised digital asset distributed over an extensive network of computers. Bitcoin was the first cryptocurrency to be operational in January 2009 after the appearance of the mysterious paper titled "Bitcoin: A Peer-to-Peer Electronic Cash System", published in 2008 under the alias "Satoshi Nakamoto" (a person or group of people) (Nakamoto 2009). The paper described a peer-to-peer payment system using electronic currency (cryptocurrencies) that could be sent directly from one party to another without using a third party (often a financial institution) to validate the transaction. The main idea behind the paper was to emulate an online shared ledger on the peer-to-peer network to validate all transactions via a blockchain, eliminating the risk of forging the register. Hence, the rise of blockchain technology provides security, privacy and a distributed ledger applicable in IoT, distributed storage systems and many more. In exchange for maintaining the blockchain (run and validate), which is energy dependent on running electronic machines, "miners" are rewarded with cryptocurrency.

The value of such a type of currency is highly volatile and speculative. It is different from any other asset on the financial market and thereby creates new possibilities for stakeholders with regard to risk management, portfolio analysis and consumer sentiment analysis (Dyhrberg 2016). Bitcoin had a market cap of over \$ 300B in December 2022. From basically nothing in January 2009 to reaching the highest price (known to date) of \$67,566.83 on November 8, 2021, Bitcoin has been highly volatile compared to other traditional forms of FIAT payment. During that

time, there have been substantial price changes over short periods. In 2017 the value of a single Bitcoin increased from \$863 on January 9, 2017, to a high of \$17,550 on December 11, 2017 (Abraham, Higdon, Nelson & Ibarra 2018). Due to its volatility and never seen behaviour like traditional currencies, Bitcoin (and cryptocurrencies in general) are extremely difficult to predict. Nevertheless, it was found that the best model for Bitcoin price volatility is the AR-CGARCH model, highlighting the significance of including both a short-run and a long-run component of the conditional variance (Katsiampa 2017). Halvor et al. find that the heterogeneous autoregressive model is suitable for Bitcoin volatility whereby trading volume further improves this volatility model (Aalborg, Molnár & de Vries 2019).

In this modern age, social media platforms allow people to share sentiments on a large scale. Hence, prompting research to study the impact of sentiment on different complex problems. Asur and Huberman attempted to solve the revenue prediction problem using tweet volume and sentiment (Asur & Huberman 2010). The same kind of data and polling results were also used by Bermingham & Smeaton to train a linear regression model to predict election results (Bermingham & Smeaton 2011). Zhang et al. leveraged sentiments on Twitter to help predict the movement of stock market indices such as Dow Jones, S&P500 and NASDAQ (Zhang, Fuehres & Gloor 2011). It was shown that significant negative emotions and opinions caused Dow to go down the next day, and lower negative emotions caused Dow to go up the next day. Consequently, studies are being conducted to correlate the speculative nature of Bitcoin prices and their sentiments on social media. Banerjee et al. studied the impact of cryptocurrency returns and Covid 19-news. A nonlinear technique of transfer entropy was used to investigate the relationship between the top 30 cryptocurrencies by market capitalisation and COVID-19 news sentiment. Results show that COVID-19 news sentiment influences cryptocurrency returns (Banerjee, Akhtaruzzaman, Dionisio, Almeida & Sensoy 2022). Zhang et al. explored the cryptocurrency market's reaction to issuers' Twitter sentiments. It was found that cryptocurrency prices react positively to Twitter sentiments (Zhang & Zhang 2022). In contrast, the trading volume reacts positively to the absolute value of Twitter sentiments in a timely manner (within a period of 24 h).

This project introduces the essential concepts of deep learning, optimisation, and sentiment analysis. We then follow up with an application in the prediction of Bitcoin's price. Chapter 2 introduces the perceptron, its development into feedforward neural networks and a derivation of the backpropagation algorithm. In Chapter 3, we looked at the different optimisation algorithms and the application of the Adam algorithm in backpropagation. Then, in Chapter 3, we reviewed the traditional and modern sentiment analysis techniques. An application for the prediction of Bitcoin price using data from different sources, sentiment analysis and LSTM NN is studied in Chapter 5. Finally, the project concludes by summarising all our results, findings, and potential improvements.

Chapter 2

Conclusion and Future Works

This study investigated how we can leverage deep learning models and sentiment analysis on tweets to predict the value of Bitcoin in the next hour.

We reviewed the basics and derived the backpropagation algorithm of feed-forward neural networks. An application of the backpropagation algorithm for feed-forward neural networks was derived. Moreover, we investigated the literature and defined different architectures of sequential neural networks used in problems involving sequential data. The literature on optimization algorithms was also reviewed. We looked at the application of the Adam optimizer in the backpropagation algorithm, the current state-of-the-art algorithm for the optimization problem. We also briefly examined sentiment analysis literature using traditional methods and the latest deep learning models involving transformers for natural language processing, roBERTa (Robustly Optimized BERT Pre-training Approach).

Finally, we defined and coded a whole pipeline from scrapping data to predicting the value of Bitcoin in the next hour using the LSTM NN model. Scrapping, cleaning and feature engineering raw data is tedious. The RMSE for the training and testing on the whole dataset were \$303.96 and \$325.96, respectively. We saw a slight decrease in performance when using only the influencers' data which might be due to the naive way we used to identify influencers: by only considering their number of followers. Furthermore, deep learning models are not easily interpretable due to their BlackBox-like architecture. It did involve a significant amount of trial.

As a continuity to this work, it would be interesting to compare our model to

traditional statistical models. Additionally, we could investigate the features to be added in the LSTM NN model and/or identify the dominant features to reduce our dataset (potentially decreasing training time).

Bibliography

Aalborg, H. A., Molnár, P. & de Vries, J. E. (2019), ‘What can explain the price, volatility and trading volume of bitcoin?’, *Finance Research Letters* **29**, 255–265.

URL: <https://www.sciencedirect.com/science/article/pii/S1544612318302058>

Abraham, J., Higdon, D. W., Nelson, J. & Ibarra, J. (2018), Cryptocurrency price prediction using tweet volumes and sentiment analysis.

Asur, S. & Huberman, B. (2010), ‘Predicting the future with social media’, *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010* **1**.

Banerjee, A. K., Akhtaruzzaman, M., Dionisio, A., Almeida, D. & Sensoy, A. (2022), ‘Nonlinear nexus between cryptocurrency returns and covid-19 news sentiment’, *Journal of Behavioral and Experimental Finance* **36**, 100747.

URL: <https://www.sciencedirect.com/science/article/pii/S2214635022000703>

Bar-Haim, R., Dinur, E., Feldman, R., Fresko, M. & Goldstein, G. (2011), Identifying and following expert investors in stock microblogs, in ‘Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing’, Association for Computational Linguistics, Edinburgh, Scotland, UK., pp. 1310–1319.

URL: <https://aclanthology.org/D11-1121>

Barbieri, F., Camacho-Collados, J., Neves, L. & Anke, L. E. (2020), ‘Tweeteval: Unified benchmark and comparative evaluation for tweet classification’, *CoRR*

abs/2010.12421.

URL: <https://arxiv.org/abs/2010.12421>

Bengio, Y., Simard, P. & Frasconi, P. (1994), ‘Learning long-term dependencies with gradient descent is difficult’, *IEEE Transactions on Neural Networks* **5**(2), 157–166.

Bermingham, A. & Smeaton, A. (2011), On using Twitter to monitor political sentiment and predict election results, *in* ‘Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2011)’, Asian Federation of Natural Language Processing, Chiang Mai, Thailand, pp. 2–10.

URL: <https://aclanthology.org/W11-3702>

Chiorrini, A., Diamantini, C., Mircoli, A. & Potena, D. (2021), Emotion and sentiment analysis of tweets using bert.

Devlin, J., Chang, M., Lee, K. & Toutanova, K. (2018), ‘BERT: pre-training of deep bidirectional transformers for language understanding’, *CoRR* **abs/1810.04805**.

URL: <http://arxiv.org/abs/1810.04805>

Dyhrberg, A. H. (2016), ‘Hedging capabilities of bitcoin. is it the virtual gold?’, *Finance Research Letters* **16**, 139–144.

URL: <https://www.sciencedirect.com/science/article/pii/S1544612315001208>

Hinton, G. & Tieleman, T. (2012), ‘Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude.’, COURSE: Neural Networks for Machine Learning, 4, 26-30.

Hochreiter, S. & Schmidhuber, J. (1997), ‘Long Short-Term Memory’, *Neural Computation* **9**(8), 1735–1780.

URL: <https://doi.org/10.1162/neco.1997.9.8.1735>

Katsiampa, P. (2017), ‘Volatility estimation for bitcoin: A comparison of garch models’, *Economics Letters* **158**, 3–6.

URL: <https://www.sciencedirect.com/science/article/pii/S0165176517302501>

- Kingma, D. P. & Ba, J. (2014), ‘Adam: A method for stochastic optimization’.
URL: <https://arxiv.org/abs/1412.6980>
- Kolen, J. F. & Kremer, S. C. (2001), *Gradient Flow in Recurrent Nets: The Difficulty of Learning LongTerm Dependencies*, pp. 237–243.
- LeCun, Y., Bengio, Y. & Hinton, G. (2015), ‘Deep learning’, *Nature* **521**(7553), 436–444.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. & Stoyanov, V. (2019), ‘Roberta: A robustly optimized BERT pretraining approach’, *CoRR* **abs/1907.11692**.
URL: <http://arxiv.org/abs/1907.11692>
- McCulloch, W. S. & Pitts, W. (1943), ‘A logical calculus of the ideas immanent in nervous activity’, *The bulletin of mathematical biophysics* **5**(4), 115–133.
- Mikolov, T., Chen, K., Corrado, G. & Dean, J. (2013), ‘Efficient estimation of word representations in vector space’.
URL: <https://arxiv.org/abs/1301.3781>
- Nakamoto, S. (2009), ‘Bitcoin: A peer-to-peer electronic cash system’, *Cryptography Mailing list at* <https://metzdowd.com> .
- Robbins, H. & Monro, S. (1951), ‘A Stochastic Approximation Method’, *The Annals of Mathematical Statistics* **22**(3), 400 – 407.
URL: <https://doi.org/10.1214/aoms/1177729586>
- Rosenblatt, F. (1958), ‘The perceptron: a probabilistic model for information storage and organization in the brain.’, *Psychol Rev* **65**(6), 386–408.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986), ‘Learning representations by back-propagating errors’, *Nature* **323**(6088), 533–536.
- Schuster, M. & Paliwal, K. (1997), ‘Bidirectional recurrent neural networks’, *IEEE Transactions on Signal Processing* **45**(11), 2673–2681.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. & Polosukhin, I. (2017), ‘Attention is all you need’.

URL: <https://arxiv.org/abs/1706.03762>

Zhang, J. & Zhang, C. (2022), ‘Do cryptocurrency markets react to issuer sentiments? evidence from twitter’, *Research in International Business and Finance* **61**, 101656.

URL: <https://www.sciencedirect.com/science/article/pii/S0275531922000447>

Zhang, X., Fuehres, H. & Gloor, P. A. (2011), ‘Predicting stock market indicators through twitter “i hope it is not as bad as i fear”’, *Procedia - Social and Behavioral Sciences* **26**, 55–62. The 2nd Collaborative Innovation Networks Conference - COINs2010.

URL: <https://www.sciencedirect.com/science/article/pii/S1877042811023895>