Informatics Institute of Technology

In Collaboration With

The University of Westminster, UK

The University of Westminster, Coat of Arms

Surpassing Time Series Forecasting Limitations using Liquid Time-stochasticity Networks

A Project Introduction by

Mr. Ammar Raneez

W1761196 | 2019163

Supervised by

Mr. Torin Wirasingha

April 2023

This is submitted in partial fulfilment of the requirements for the BSc (Hons) Computer Science degree at the University of Westminster.

Contents

List of Tables

List of Abbreviations

AI	Artificial Intelligence.
API	Application Programming Interface.
ARIMA	Autoregressive Integrated Moving Average.
BPTT	Back-Propagation Through Time.
BTC	Bitcoin.
CT-GRU	Continuous-time Gated Recurrent Unit.
CT-RNN	Continuous-time Recurrent Neural Network.
DL	Deep Learning.
GPU	Graphics Processing Unit.
LSTM	Long Short-Term Memory.
LTC	Liquid Time-constant.
LTS	Liquid Time-stochasticity.
ML	Machine Learning.
(s)MAPE	Symmetric Mean Absolute Product Error.
MASE	Mean Absolute Scaled Error.
MSE	Mean Squared Error.
N-BEATS	Neural Basis Expansion Analysis for interpretable Time Series
NLP	Natural Language Processing.
ODE	Ordinary Differential Equations.
POC	Proof-Of-Concept.
REST	Representational State Transfer.
RMSE	Root Mean Squared Error.
RNN	Recurrent Neural Network.
TS	Time Series.
SDE	Stochastic Differential Equations.

1. CHAPTER OVERVIEW

In this chapter, the author aims to identify and provide the reader with an overview of the current issues in time series forecasting and highlight what a liquid time-stochasticity network is and what it aims to solve. In detail, the author will define the problem and evaluate the necessary literature to arrive at a justifiable research gap, respective research objectives, and challenges that would arise. The novelty within the chosen problem and the proposed solution are also stated.

2. PROBLEM DOMAIN

2.1 Time series forecasting

TS forecasting is a significant business issue and an area where ML could create an impact. It is the foundation for contemporary business practices, including pivotal domains like customer management, inventory control, marketing, and finance. As a result, it has a comprehensive financial impact, with millions of dollars for subtle improvements in forecasting accuracy (Jain, 2017).

Having said that, although ML and DL have outperformed classical approaches for NLP and computer vision, the domain of TS still seems to be a struggle compared to classical statistical methodologies (Makridakis et al., 2018a;b). Most of the top-ranking methods in the M4 competition were ensembles of traditional statistical techniques (Makridakis et al., 2018b), while regular ML methods were nowhere near.

Therefore, this competition's winner was a hybrid model of LSTM and classical statistics (Smyl, 2020). Furthermore, they claimed that the only way to improve the accuracy of TS forecasting was by creating such hybrid models, which the author aspires to challenge in this research project.

2.2 Liquid Time-Constant (LTC) networks

LTCs are neural ODEs: hidden layers are not specified; instead, the derivative of hidden states is parameterized by neural networks (Chen et al., 2019). RNNs are successful algorithms for TS data modelling, if there exist continuous time-hidden states determined by ODEs (Chen et al., 2019). Studies show that existing algorithms such as the CT-RNN (Funahashi and Nakamura, 1993;

Rubanova, Chen and Duvenaud, 2019) and CT-GRU (Mozer, Kazakov and Lindsey, 2017) produce such performance. However, they have issues with expressivity and fixed behaviour once trained (Hasani et al., 2020). Therefore, the question arises: what would happen if there were unexpected changes to the characteristic of the inputs during inference? Additionally, these algorithms lose in generalization compared to even a simple LSTM network (Hasani et al., 2021), which raises another question, what is the point of defining a different and *'fancy'* approach if they cannot work well in real-world applications?

Hasani et al. state that LTCs can *"identify specialized dynamical systems for input features arriving at each time point"* (2020, p1). The ability to exhibit stable and bounded behaviour demonstrates that the proposed approach yields better expressivity than traditional implementations. The LTC state and their respective time constant "*exhibit bounded dynamics and ensure the stability of the output dynamics*", which is a prominent factor when inputs increase relentlessly (Hasani et al., 2020).

However, it is important to note that the underlying concepts used within the architecture is obsolete (Duvenaud, 2021) and lacks adaptability to noisy data and data with high volatility. The LTS algorithm presented by the author is an enhancement to this LTC considering these limitations.

2.3 Cryptocurrencies

The word 'crypto' has been an enormous buzzword recently, especially BTC. It has even come to the point where crypto and BTC are used interchangeably.

Cryptocurrencies are a fully decentralized digital currency form; it is a form of a peer-to-peer system without the need for a third party, thus enabling safer online transactions (S. Nakamoto, 2008). BTC is the first and the most popular digital currency to date, piquing many academic researchers' interest (Rahouti et al., 2018).

However, being at the forefront of the digital currency world, it has developed high volatility, making it difficult to predict future rates and a disadvantage for multiple investors (Kervanci and Akay, 2020). Despite that, recent advances in ML and statistics have shown acceptable results in the analysis and prediction of cryptocurrencies, yet the root cause of these algorithms persists: they are static.

3. PROBLEM DEFINITION

Most applications of ML available do perform quite well (ex: image classification, transfer learning, NLP), yet, as mentioned, the field of TS forecasting is subpar (Makridakis et al., 2018a;b). Existing TS forecasting algorithms cannot adapt to unforeseen changes in data streams (Hasani et al., 2020); consequently, they perform relatively poorly when applied to TS forecasting.

It is worth noting that the neural ODEs (Chen et al., 2019), and notably the LTC architecture proposed by Hasani et al. (2020) rectify this to some extent; however, their proposed solution does not have the ability to model instantaneous changes. Therefore, they perform poorly when applied to areas of high volatility (Duvenaud, 2021) (in this case: the forecasting of BTC).

3.1 Problem statement

Existing TS forecasting systems cannot adapt to incoming data streams with unexpected changes and characteristics that are different from the data on which they were trained.

4. RESEARCH MOTIVATION

The field of AI, particularly neural networks, has been growing exponentially recently, alongside intriguing research. However, as mentioned by Hasani et al. (2020), the issue of networks being static and unable to adapt to varying characteristics could prove to be a limitation for the future of intelligent systems, TS in particular. Therefore, this research is expected to facilitate further exploration by trying to aid in driving the domain forward.

5. RESEARCH GAP

The literature defines only a single paper for the LTC architecture (Hasani et al., 2020). Where every other work is related to the family of neural ODEs (CT-RNN (Rubanova, Chen and Duvenaud, 2019) and CT-GRU (Mozer, Kazakov and Lindsey, 2017)) and the secondary problem domain of cryptocurrencies and TS.

Gap in existing forecasting algorithms

It is also worth noting that, because of this, existing forecasting solutions are all implemented using traditional deep neural net approaches (ex: LSTM (Hochreiter and Schmidhuber, 1997)) that are static and hence have the limitation of not being able to learn and adapt during inference (Hasani et al., 2020), which results in the model's accuracy degrading overtime – a '*data drift*' (Poulopoulos, 2021).

Gap in LTC

The proposed LTC architecture uses a sequence of linear ODEs, which are now considered obsolete and lack rapid adaptability (Duvenaud, 2021). Recent advancements in this field suggest the usage of SDEs instead, as they are more flexible. An additional issue is that ODEs model 'deterministic dynamics' – uncertainty, or any unobserved interactions cannot be modelled, which is inevitable in TS data. Therefore, the way forward is to build the LTC architecture with SDEs instead. As implied, the algorithm proposed is the author's own, giving it the name of "Liquid Time-Stochasticity", as it is no longer constant.

Gap in bitcoin forecasting solutions

The work available on BTC forecasting has not considered exogenous factors that could have an impact (Roy et al., 2018; Rizwan et al., 2019; Fleisher et al., 2022). Therefore, a significant concern is that they cannot adapt well; in other words, they are not robust. Factors that could influence the price are as follows:

- Tweet sentiment & volume
- Google Trends

Cryptocurrency forecasting is as reliable as palm reading since it is an open system. Hence, uncontrollable factors such as a country's law can affect the price. Despite this, researchers can consider certain factors, such as the ones mentioned above. Abraham et al. (2018) identified that the above factors correlate with the price of BTC; therefore, it is vital to consider them when building such systems in the future.

6. CONTRIBUTION TO THE BODY OF KNOWLEDGE

In a nutshell, the author desires to answer the following question:

- Would a novel architecture built by the LTS algorithm be an advancement for TS forecasting?

6.1 Research domain contribution

An implementation of the LTS algorithm will be developed, following the proposed architecture, to facilitate the model creation. Additionally, the algorithm built will be generalized without being problem-specific so that researchers can apply it elsewhere to evaluate its performance and identify whether it would also be an advancement to those domains.

Having understood the issues in the current literature, a solution is a dynamic algorithmic architecture that can adapt and change its underlying mathematical expressions and evaluation strategies based on changes in the incoming data streams. Additionally, further enhancements are required to avoid sudden and tiny changes that are common in TS data. Being able to adapt to unforeseen changes and being highly expressive could be the stepping stone within the TS forecasting community to aid in future research.

6.2 Problem domain contribution

Based on the above critique, creating a more robust forecasting solution considering the mentioned factors (Twitter, Google Trends) could mean that the highly volatile market of cryptocurrencies could be predicted much more efficiently and be the way forward for investors.

Furthermore, there is no direct data source that the author can utilize for building the model. Therefore, dedicated scripts are developed to obtain the said data. These scripts will be available for public access to support future research on BTC forecasting.

7. RESEARCH CHALLENGE

Existing architectures scale up, and the LTC scales down - with more expressive nodes (Hasani et al., 2020). Having adapted to the “deeper is usually better” mindset of deep neural nets, a challenge opens up in **identifying the requirement of scaling down and what a neural ODE aims to solve**.

LTCs are a new approach with only a single research paper regarding its proposed solution. Currently, it is only in the experimental stage and utilizes a novel formulation compared to other existing neural ODEs (Hasani et al., 2020).

The broader domain of neural ODEs (Chen et al., 2019) **is also relatively new**; hence the scarcity of references could create more challenges for further research or implementation of systems.

SDEs are an advanced topic in mathematics, and modelling them as neural SDEs have had a couple of research conducted; however, they were primarily for specific purposes. Therefore, **no generic papers exist for neural SDEs, unlike neural ODEs, which makes modelling difficult**.

Currently, existing TS forecasting systems are built using statistical ensemble methods (Makridakis et al., 2018b) or traditional neural net architectures (Hasani et al., 2021), which creates a new challenge where **there is no reference implementation**.

The chosen domain of application is an open system. **Open system forecasting is usually poor and generally difficult to beat the naïve forecast** (A naive forecast is not necessarily bad, 2014) since it can depend on any external factor. Therefore, there is the possibility of discouragement from continuing the research if the results are not as expected.

8 RESEARCH QUESTIONS

RQ1: What are the recent advancements in TS algorithms that can be considered when building the LTS?

RQ2: How can the algorithm be used to implement a TS forecasting system, and how will the challenges faced be overcome?

RQ3: What contributions can be made to the chosen domain?

9. RESEARCH AIM

The aim of this research is to design, develop & evaluate a novel algorithm (LTS) inspired by the LTC so that it can build intelligent systems by developing a novel architecture for TS forecasting, which could be the stepping stone to be further expanded to other domains as well.

Specifically, this research project will produce a TS forecasting system utilizing the LTS algorithm, focused on the forecasting of BTC.

10. RESEARCH OBJECTIVES

Research objectives are milestones that the author must achieve for the research to succeed.

Table 1: Research Objectives (*Self-Composed*)

Objective	Description	Learning Outcomes	Research Questions
Problem Identification	Understand and document the identified problem and provide reasoning on what makes it novel. **R**O1: Conduct research on a domain of interest and identify a comprehensive enough issue that requires solving. **R**O2: Delve deeper into the identified problem to obtain a general understanding on how to approach solving the problem. **R**O3: Split the problem down into manageable subsections so it is easier to digest and to solve one section at a time. **R**O4: Design a respective schedule, associated deliverables, and the Gantt chart.	LO1, LO2	**R**Q1
Literature Review	Collate relevant information by reading, understanding, and evaluating previous work. **R**O5: Conduct preliminary studies and investigations on existing TS forecasting systems and algorithms. **R**O6: Analyze the requirement for specialized TS algorithms. **R**O7: Conduct research on neural ODEs, LTCs & SDEs. **R**O8: Obtain deep insights into the architecture behind the LTC.	LO1, LO4, LO5	**R**Q1

	RO9: Research and obtain insights on factors affecting the price of BTC. **R**O10: Research on existing BTC forecasting and related open market systems. **R**O11: Research on necessary ML techniques and evaluation approaches.		
Requirement Elicitation	Collect and analyze project requirements using appropriate tools and techniques. **R**O12: Analyze stakeholders and understand their viewpoints and concerns. **R**O13: Gather the requirements and architectures of LTCs and SDEs. **R**O14: Collate the most up-to-date details of BTC and obtain insights on the perspectives of the end users. **R**O15: Design necessary diagrams to justify the product's specification.	LO1, LO3	**R**Q1
Design	Design the architecture and a corresponding system capable of effectively solving the identified problems. **R**O16: Design necessary diagrams required to understand the algorithm RO17: Design diagrams required to understand the supplementary system being developed. **R**O18: Design the novel algorithm and analyze its complexities.	LO1	**R**Q2, **R**Q3
Implementation	Implement a system that is capable of addressing the research gaps. **R**O19: Design, evaluate and pick necessary technologies best-suited for the implementation. **R**O20: Develop an efficient LTC implementation.	LO1, LO5, LO6, LO7	**R**Q2, **R**Q3

	RO21: Build on the LTC to implement the LTS. **R**O22: Integrate the algorithm developed into a TS forecasting application. **R**O23: Integrate the intelligent system into a client application to display forecasts. **R**O24: Design and implement an automated flow to update the built network with the latest data. **R**O25: Design and implement a pipeline for easy deployments. **R**O26: Consider any legal, social, ethical & professional issues upon implementation.		
Evaluation	Effectively test the algorithm implemented, the system, and the respective data science model using recommended techniques. **R**O27: Evaluate the developed algorithm and the respective model against the evaluation metrics researched in the literature review. **R**O28: Create a test plan & test cases and perform unit, performance, and integration testing.	LO4	**R**Q2, **R**Q3
Documentation	Document the progression of the research project and inform about any challenges faced. **R**O29: Create a coherent report of new skills obtained, evaluations, contributions etc., and ensure that all the above-stated objectives are met.	LO6, LO8	**R**Q3

11. CHAPTER SUMMARY

In this chapter, the author provided an overview of the research project carried out, respective reasons for the research and problem to be a novelty, and the challenges they can face upon solving it. Furthermore, the necessary goals that must be aimed to consider the research successful were proposed and mapped to the learning outcomes that must be attained by the chosen degree.

REFERENCES

S. Nakamoto, (2020). *Bitcoin: A peer-to-peer electronic cash system*. Available from https://bitcoin.org/bitcoin.pdf [Accessed 25 September 2022].

Rahouti, M., Xiong, K. and Ghani, N. (2018). Bitcoin Concepts, Threats, and Machine-Learning Security Solutions. *IEEE Access*, 6, 67189–67205. Available from https://doi.org/10.1109/ACCESS.2018.2874539 [Accessed 25 September 2022].

Kervanci, I. sibel and Akay, F. (2020). Review on Bitcoin Price Prediction Using Machine Learning and Statistical Methods. *Sakarya University Journal of Computer and Information Sciences*. Available from https://doi.org/10.35377/saucis.03.03.774276 [Accessed 25 September 2022].

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018a). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13 (3), e0194889. Available from https://doi.org/10.1371/journal.pone.0194889 [Accessed 25 September 2022].

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018b). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34 (4), 802–808. Available from https://doi.org/10.1016/j.ijforecast.2018.06.001 [Accessed 25 September 2022].

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36 (1), 75–85. Available from https://doi.org/10.1016/j.ijforecast.2019.03.017 [Accessed 25 September 2022].

Hasani, R. et al. (2020). Liquid Time-constant Networks. Available from https://doi.org/10.48550/arXiv.2006.04439 [Accessed 25 September 2022].

Chen, R.T.Q. et al. (2019). Neural Ordinary Differential Equations. Available from https://doi.org/10.48550/arXiv.1806.07366 [Accessed 25 September 2022].

Poulopoulos, D. (2021). Is “Liquid” ML the answer to autonomous driving? *Medium*. Available from https://towardsdatascience.com/is-liquid-ml-the-answer-to-autonomous-driving-bf2e899a9065 [Accessed 25 September 2022].

Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9 (8), 1735–1780. Available from https://doi.org/10.1162/neco.1997.9.8.1735 [Accessed 25 September 2022].

BI4ALL. (2021). Supervised Machine Learning in Time Series Forecasting. *BI4ALL - Turning Data Into Insights*. Available from https://www.bi4all.pt/en/news/en-blog/supervised-machine-learning-in-time-series-forecasting/ [Accessed 12 October 2022].

Hasani, R. et al. (2021). Liquid Neural Networks. *YouTube*. Available from https://www.youtube.com/watch?v=IlliqYiRhMU&t=350s. [Accessed on 30 Sep. 2022].

Mozer, M.C., Kazakov, D. and Lindsey, R.V. (2017). Discrete Event, Continuous Time RNNs. Available from https://doi.org/10.48550/ARXIV.1710.04110 [Accessed 14 October 2022].

Funahashi, K. and Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6 (6), 801–806. Available from https://doi.org/10.1016/S0893-6080(05)80125-X [Accessed 14 October 2022].

A naive forecast is not necessarily bad. (2014). *The Business Forecasting Deal*. Available from https://blogs.sas.com/content/forecasting/2014/04/30/a-naive-forecast-is-not-necessarily-bad/ [Accessed 15 October 2022].

Rubanova, Y., Chen, R.T.Q. and Duvenaud, D. (2019). Latent ODEs for Irregularly-Sampled Time Series. Available from https://doi.org/10.48550/ARXIV.1907.03907 [Accessed 18 October 2022].

Duvenaud, D (2021). Directions in ML: Latent Stochastic Differential Equations: An Unexplored Model Class. *YouTube*. Available from https://www.youtube.com/watch?v=6iEjF08xgBg. [Accessed on 30 Sep. 2022].

Chaman L. Jain. Answers to your forecasting questions. Journal of Business Forecasting, 36, Spring 2017.

Abraham, J., Higdon, D., Nelson, J. and Ibarra, J. (2018). Cryptocurrency Price Prediction Using Tweet Volumes and Sentiment Analysis. *SMU Data Science Review:* Vol. 1: No. 3, Article 1. Available at: https://scholar.smu.edu/datasciencereview/vol1/iss3/1

Roy, S., Nanjiba, S. and Chakrabarty, A. (2018). Bitcoin Price Forecasting Using Time Series Analysis. *2018 21st International Conference of Computer and Information Technology (ICCIT)*.

December 2018. Dhaka, Bangladesh: IEEE, 1–5. Available from https://doi.org/10.1109/ICCITECHN.2018.8631923 [Accessed 25 September 2022].

Rizwan, M., Narejo, S. and Javed, M. (2019). Bitcoin price prediction using Deep Learning Algorithm. *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*. December 2019. Karachi, Pakistan: IEEE, 1–7. Available from https://doi.org/10.1109/MACS48846.2019.9024772 [Accessed 26 September 2022].

Fleischer, J.P. et al. (2022). Time Series Analysis of Cryptocurrency Prices Using Long Short-Term Memory. *Algorithms*, 15 (7), 230. Available from https://doi.org/10.3390/a15070230 [Accessed 26 September 2022].