Informatics Institute of Technology

In Collaboration With

University of Westminster, UK

University of Westminster, Coat of Arms

A Novel Approach to Time Series Forecasting using Liquid Time-constant Networks

A Project Proposal by

Mr. Ammar Raneez

W1761196 | 2019163

Supervised by

Mr. Torin Wirasingha

September 2022

This Project Proposal 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 Figures

Acronyms

AI Artificial Intelligence.

BPTT Back-Propagation Through Time.

BTC Bitcoin.

CTN-GRU Continuous-time Gated Recurrent Unit.

CT-RNN Continuous-time Recurrent Neural Network.

DL Deep Learning.

GUI Graphical User Interface.

LSTM Long Short-Term Memory.

LTC Liquid Time-constant.

ML Machine Learning.

MSE Mean Squared Error.

NLP Natural Language Processing.

ODE Ordinary Differential Equations.

RMSE Root Mean Squared Error.

RNN Recurrent Neural Network.

TS Time Series.

1. Introduction

In this research project, the author aims to identify the current issues with TS forecasting by researching existing literature and attempts a method of rectification by a novel approach utilizing a modern neural net architecture. The architecture put forward will attempt to transform the current traditional approaches of TS analysis which could be enhanced in future to create more intelligent systems, alongside an example implementation of a forecasting model for BTC,

This document defines the problem, the research gap, and the research challenge, alongside the necessary proof required for justification.

2. Problem Domain

2.1 Time Series Forecasting

TS forecasting is a significant business issue and an area where ML could create a promising impact. It serves as the foundation for contemporary business practices, including pivotal domains like customer management and inventory control, marketing and finance. As a result, it has a comprehensive financial impact, within millions of dollars for each additional point of forecasting accuracy (Jain, 2017; Kahn, 2003).

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 point of struggle when compared to classical statistical methodologies (Makridakis et al., 2018a;b). For instance, out of a total of 60 submissions, the six "pure" ML methods submitted to the M4 competition were ranked 23, 37, 38, 48, 54, and 57, and the majority of the top-ranking methods were ensembles of traditional statistical techniques (Makridakis et al., 2018b).

It is therefore worth mentioning that the winner of this competition was a hybrid model of an LSTM (Smyl, 2020), which went on to claim that the only way forward for improving TS forecasting accuracy was by creating hybrid models, which the author aspires to challenge on this research project.

2.2 Liquid Time-Constant Networks

RNNs with continuous time hidden states determined by ODEs are effective algorithms for TS data modelling (Chen et al., 2018). Studies show that existing algorithms such as the CT-RNN and CT-GRU produce such performance, however, have issues in expressivity and a 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?

"*LTCs can identify specialized dynamical systems for input features arriving at each time point.*" The ability to exhibit stable and bounded behaviour demonstrates that the proposed approach yields more predominant expressivity (Hasani et al., 2020).

LTCs state and their respective time-constant exhibit bounded dynamics and assure the stability of the output dynamics, which is a prominent factor when inputs increase relentlessly (Hasani et al., 2020).

2.3 Cryptocurrencies

The word "Crypto" is a huge buzzword in recent times, BTC, especially. It has even come to the point where Crypto and BTC are used interchangeably.

Cryptocurrencies are a form of digital currency that is fully decentralized (Rahouti et al., 2018); it's a form of peer-to-peer system without the need for a third party, thereby enabling safer online transactions (S. Nakamoto, 2008). In the world of digital currencies, BTC is the first and the most popular cryptocurrency to date, which has piqued the interest of many academic researchers (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 somewhat promising results in the analysis and prediction of cryptocurrencies, yet, the root cause of these Algorithms persists: they are static.

3. Problem Definition

As of writing this report, there is no application of liquid time-constant networks in any domain since this field of neural networks has only recently been announced. Existing intelligent systems utilize more traditional approaches of neural nets developed some time ago.

Having mentioned that, most applications of ML available do perform quite well (Ex: Image classification, Transfer learning, NLP etc.), yet, as mentioned, the field of TS forecasting seems to be subpar. Existing TS forecasting algorithms cannot adapt to unforeseen changes in data streams and perform quite poorly when used in areas of high volatility (In this case: the forecasting of BTC).

To aid with further research on this new concept of ODE neural networks, it is identified that the building of an LTC and its application on an ML domain that still can struggle could be the stepping stone for future intelligent systems – and as a supplement, provide hope to crypto investors for easier predictions.

3.1 Problem Statement

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

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. This research, therefore, is expected to facilitate further exploration by attempting to aid in driving the domain of TS forward.

5. Related Work

Since there is no existing work on LTCs, the author will break down work towards general TS forecasting and its application in BTC forecasting.

Table 1: Related Work

Citation	Summary	Technique Used	Contributions	Limitations
TS Forecasting (General)				
(Hochreiter and Schmidhuber, 1997)	An algorithm that learns to bridge minimal time lags by enforcing constant error flows. It learns	LSTM. A recurrent network architecture with an appropriate gradient-based learning algorithm	Improved performance for short sequence predictions. Overcame error back-flow problems present	Prediction capacity limits long sequence performance, where the MSE and RMSE rise unacceptably

	much faster, creates more successful runs and has the capability to solve complex tasks that have not been solved before.		in conventional BPTT, where they tended to blow up or vanish.	high. Therefore, is not an ideal solution for predictions of the distant future.
(“Autoregressive Integrated Moving Average (ARIMA)”, 2021)	A statistical analysis model to understand the dataset or predict future trends. This model depends on the past values to predict the future and uses lagged moving averages to smoothen the data.	ARIMA. A model that predicts future behaviour based on past behaviour	Improved performance for TS forecasting data that correlate with values ahead of time and before.	Does not handle well with nonlinear data and long-term forecasting. Further, it performs best on univariate analysis.
(Oreshkin et al., 2020)	An architecture that solves the univariate time series point forecasting problem. It carries some benefits some of which are being	N-Beats. A deep neural net architecture based on backward and forward residual links.	Outperformed the M4 competition’s winner and improved statistical benchmark forecast accuracy.	Tailored specifically for univariate TS analysis, therefore, would not perform well on multivariate analysis.

	understandable, easily applicable to multiple other fields and being fast to train.			
BTC Forecasting				
(Roy et al., 2018)	Applied statistical analysis to predict the price of BTC using data from 2013 to 2017. Applied the ARIMA model and obtained an overall accuracy of 90% for deciding weighted costs volatility.	ARIMA	Improved overall insights obtained and added context to future predictions based on past values, alongside scoring an overall lower RMSE than other ML solutions.	Trained on data only between 2013 and 2017, and is capable of forecasting for ten consecutive days.
(Rizwan et al., 2019)	Compared the usage of LSTM and ARIMA models for the prediction of BTC, however found out that these models aren't very efficient. Used	GRU	Improved existing models built using RNN and LSTM by producing better accuracy and lower MSE, alongside taking much less time to train.	Lack of updating solution against the latest available data.

	GRU and eventually gained a higher overall accuracy.			
(Fleischer et al., 2022)	Focused on the volatility and understanding the behaviour of cryptocurrencies. Trained an LSTM model using BTC close price values to predict future prices.	LSTM	Beat performance of ARIMA on longer runtime training.	Limited to univariate and does not consider other input params (Ex: high, low, volume), and is capable of forecasting only one day.

6. Research Gap

The literature defines only a single paper for the proposed algorithmic solution - where every other piece of work is not directly related to the algorithm - but is to the secondary problem domain of Cryptocurrencies and TS. In addition, no algorithmic solution exists for the proposed LTC architecture for model implementation.

It is also worth noting that, because of this, existing forecasting solutions are all implemented using traditional approaches (Ex: LSTM ((Hochreiter and Schmidhuber, 1997)) of neural nets 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).

7. Research Contribution

In summary, the author's contributions are as follows:

- **Time Series Forecasting**: A novel implementation that can learn during inference that could also expand to other domains.
- **LTC**: Implementation of the LTC algorithm for model creation.

7.1 Technological Contribution

An implementation of the LTC algorithm will be developed, following the architecture proposed, to facilitate the model creation. It is hypothesized that the algorithm built will be generalized without being problem-specific so that it could be applied elsewhere - to evaluate its performance and identify whether the LTC would also be an advancement to other domains.

As a byproduct of the primary research, the author will attempt to explore and utilize this developed algorithm of the LTC in a way to forecast BTC – which, as identified, has not been attempted.

7.2 Domain Contribution

Having understood the issues in current literature, it is likely that a solution capable of solving the mentioned issues could be an advancement for future research. Being able to adapt to unforeseen changes and being highly expressive could mean that the highly volatile market of Cryptocurrencies would be able to be predicted much more efficiently and be the way forward for investors.

In addition, it is hypothesized to be an advancement for TS forecasting by identifying whether the newly developed LTC proposed does provide strong robustness and accuracy and outperforms currently existing TS forecasting approaches. Or whether it could be enhanced to be used in other domains altogether.

8. Research Challenge

LTCs are a new domain with limited research and only a single research paper regarding its proposed solution. Therefore, it is critical to identify what problem it intends to solve. Currently, it is only in the experimental stage and is a new approach to ODEs, hence requires the identification of limitations in existing ODEs.

Existing architectures scale up, and the LTC scales down - with more expressive nodes. Having adapted to the “deeper is usually better” mindset where architectures of several layers reside, a challenge opens up in identifying the requirement for going in the opposite direction (of scaling down).

Currently, existing TS forecasting systems are built using Ensemble statistical methods or traditional neural net architectures. This creates a new challenge where the architecture proposed has not been implemented yet in a way that could be used in intelligent systems.

The scarcity of references could henceforth create more challenges for further research or implementation of systems.

9. Research Questions

RQ1: What are the recent advancements in TS forecasting systems that can be considered when building the LTC algorithm?

RQ2: What will the implemented algorithm contribute to TS forecasting, and could it expand to other domains?

RQ3: How can the system facilitate the predictions of crypto for investors?

10. Research Aim

The aim of this research is to design, develop & evaluate the LTC algorithm in a way such that it is capable of building intelligent systems by developing a novel approach to TS forecasting, which could be the stepping stones to be further expanded to other domains as well.

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

The researched knowledge will be put forward and the hypothesis of whether the LTC algorithm can be applied to the selected domain will be evaluated.

11. Research Objectives

The accomplishment of the ensuing research objectives is anticipated to meet the aims and provide answers to the research questions listed above. These goals represent milestones that must be achieved for the research to be considered successful.

Table 2: Research Objectives

Objective	Description	Learning Outcomes
Literature Review	Collate relevant information by reading, understanding and evaluating previous work • **R**01: Conduct preliminary studies and investigations on existing TS forecasting systems. • R02: Analyze the requirement for specialized TS algorithms. • **R**03: Conduct a preliminary study on LTCs. • R04: Obtain deep insights in the architecture behind the LTC.	LO2, LO4, LO5
Requirement Analysis	Collect and analyze project requirements, using appropriate tools and techniques • **R**O1: Gather requirements and architecture of the LTC. • **R**O2: Collate the most up-to-date details of BTC. • **R**O2: Get insights from technology and domain experts.	LO1, LO2, LO3
Design	Design the architecture and a corresponding system capable of effectively solving the identified problems. • **R**O1: Design an efficient approach for the LTC algorithm. • **R**O2: Design an automated flow to update the built network with the latest data. • **R**O3: Design an ML pipeline for easy deployments.	LO1
Implementation	Implement a system that is capable of addressing the mentioned research gaps.	LO1, LO5, LO6, LO7

	• **R**O1: Implement the LTC algorithm in a way capable of model building. • **R**O2: Integrate the developed algorithm into a TS forecasting application. • **R**O3: Integrate the developed intelligent system into the prototype to display forecasts.	
Evaluation	Effectively test the implemented algorithm, the system, and the respective Data Science model using recommended techniques. • **R**O1: Create a test plan & test cases and perform unit, performance and integration testing. • R02: Evaluate the developed algorithm and the respective model against the mentioned benchmarking metrics.	LO4
Documentation	Document progression of the research project and notify of any faced challenges.	LO6, LO8

12. Project Scope

Concerning the granted time for this research project, the scope is as follows.

12.1 In-scope

- Implementation of the LTC algorithm capable of being used like currently existing solutions (Ex: RNN's).
- A system capable of forecasting the rate of BTC.
- Creation of a system utilizing the mentioned algorithm.
- Evaluation and comparison of the implemented system against currently existing solutions to validate or invalidate the hypothesis.
- A GUI capable of forecasting for multiple days.

12.2. Out-scope

- Application of the implemented algorithm in other domains to further justify the hypothesis.
- Forecast multiple different cryptocurrencies.
- Ability to take other external factors into consideration, such as social media, legislation and laws, and country advertisements for handling digital currency.

12.3 Desirables

- Evaluate implementation against the M4 competition to validate the future of TS forecasting algorithms.
- Incremental learning – The model must be updated and trained with the latest data automatically (Another existing research gap).
- Consider twitter volume as an external hyperparameter (Another existing research gap).

12.4 Prototype Diagram

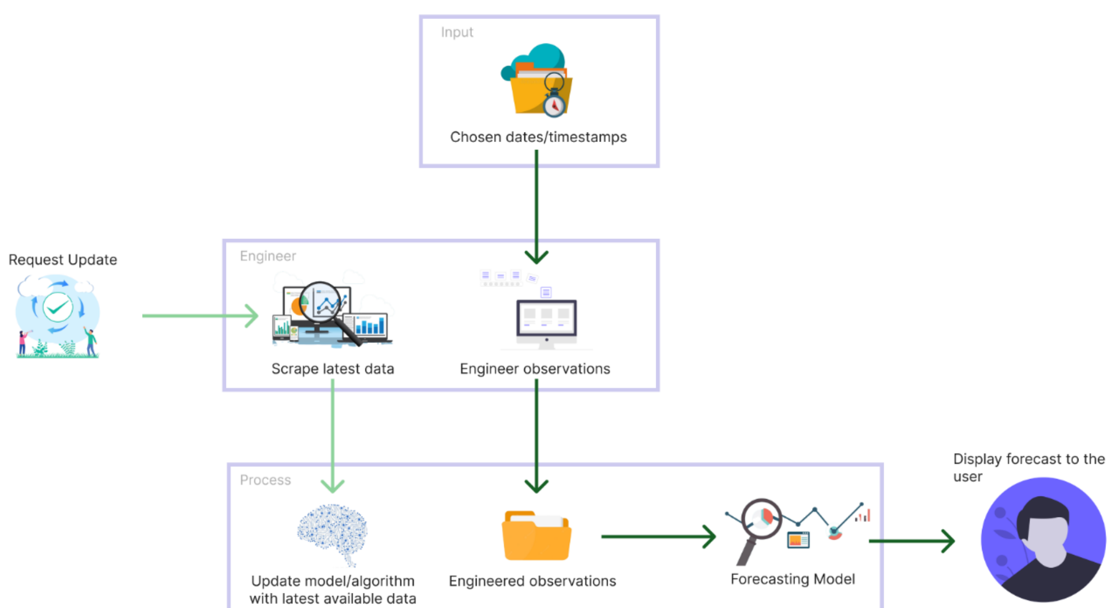

Figure 1: Prototype Feature Diagram (*Self-Composed*)

13. Proposed Methodology

13.1 Research Methodology

Methodologies suitable for the research project have been evaluated and chosen from the predefined Saunders Research Onion Model (Saunders, Lewis, and Thornhill, 2007).

Table 3: Research Methodology

Philosophy	The **Pragmatism** philosophy was chosen since the research question is of paramount importance. Additionally, as the outcome of this research, it is expected to validate/invalidate the developed hypothesis alongside necessary benchmarking comparisons.
Approach	The **deductive** approach was chosen over the inductive since the final analysis and evaluation will be quantitative that aims to deduce the hypothesis.
Strategy	**Archival Research** and **Action Research** were chosen as the strategy of data collection. The first since the research topic is more modern, hence the principal source of data collection would be research documents. Action Research will also be included since the development process will likely be an iterative approach of diagnosis, planning, taking action & evaluation.
Choice	**Multi-method** will suit the proposed research project most since qualitative analysis would be a suitable supplement to the primary quantitative approach, however will not be used as a combination.
Time Horizon	The **Longitudinal** time horizon was chosen over the cross-sectional time horizon since the developed model would need to be updated & documented frequently, tested and evaluated with the latest available data.
Techniques and procedures	As a form of **Data Collection & Analysis**, as many sources as possible will be used since there are finite resources. Statistics, reports, journals, articles and observations will be the primary mediums.

13.2 Development Methodology

13.2.1 Life Cycle Model

Agile was chosen as the research development Life Cycle since heavy iterative development is required.

13.2.2 Design Methodology

Object-Oriented Analysis & Design (OOAD) was chosen as the Design Methodology since it supports increments and extensions with reusability.

13.2.3 Evaluation Methodology

Based on research, the most suitable method for evaluating TS forecasting systems is a Walk-forward validation. This is identified to be a realistic way of evaluation since the model must be updated once new data is available (Falessi et al., 2020).

Benchmarking

MAE and RMSE will be used to benchmark the system to produce adequate comparisons against existing solutions and validate/invalidate the hypothesis.

13.3 Project Management Methodology

A combination of PRINCE2 and Agile will be followed by the author. The project will require many iterations and improvements since the implementation is novel and there exists no reference. Alongside multiple iterations its best implemented being divided into multiple chunks and focusing on each chunk at a time with a plan-based approach.

13.3.1 Schedule

Gantt Chart

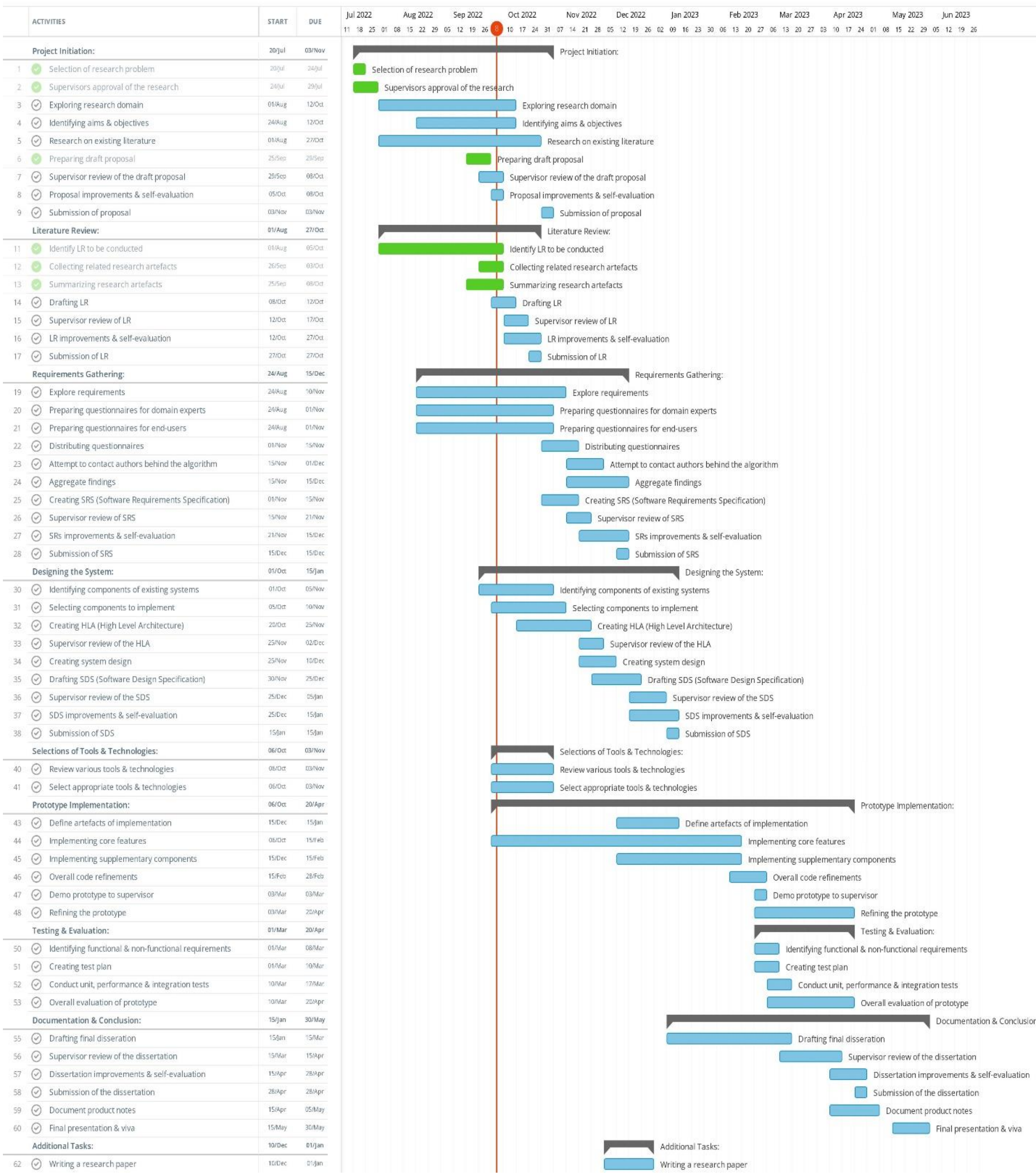

Figure 2: Gantt Chart (*Self-Composed*)

Deliverables

Table 4: Deliverables & Dates

Deliverable	**Date**
Literature Review	27th October 2022

Critical analysis of related work & solutions.	
Project Proposal & Ethics Forms Initial proposal of the research to be conducted.	3rd November 2022
Software Requirement Specification The document that defines the requirements that must be met to prototype and collect data.	24th November 2022
Proof Of Concept & Implementation Presentation An initial Implementation of the proposed system.	23rd December 2022
Project Specifications Design & Prototype A prototype of the system with the core features and am accompanying document specifying the design followed & an overview of the implemented algorithm.	23rd February 2023
Test & Evaluation Report Documentation of test finding and evaluations conducted on the prototype.	23rd May 2023
Draft Project Report A draft submission of the final Thesis to get evaluations.	30th May 2023
Final Thesis Final submission of the thesis with complete documentation of the project's journey.	27th April 2023

13.3.2 Resource Requirements

Software Requirements

- **Operating System (Windows/Linux/macOS)** – Windows will be the default since it provides easy access to required development environments and tools.

- **Python** – Will be used to create the neural net & the respective model. Python is an all-purpose language with multiple tools specifically designed for data science.
- **TensorFlow | Scikit-learn** – Provides libraries that facilitate data science in Python.
- **Node** – For seamless communication and integration between the client and the model.
- **React | NextJS** – To develop the client side of the application. A fast performant library is required to prevent lags and other performance issues.
- **VSCode | PyCharm** – Environment to facilitate application development.
- **Google Colab | Jupyter Notebook** – Development environment for building the forecasting model.
- **Zotero** – Manage references and research artefacts.
- **Overleaf | MS Office | GSuite | Figma | Canva | Draw.io** – Tools to create reports, figures, diagrams & documents and backup artefacts.
- **GitHub** – Track, version & manage development code.

Hardware Requirements

- **Core i5 Processor (8th gen) or above** – for long-running intensive workloads.
- **8GB Ram or above** – to manage model training, multiple development environments & multitasking.
- **Disk space of approx. 20GB** – to store application code & data.

Data Requirements

- **BTC price observations** – from Bitfinex.

Skill Requirements

- Creation of TS forecasting systems.
- Implementation of a raw neural net.
- Ability to create optimized & scalable DL models.
- Ability to develop optimized client-side charts & user interfaces that dynamically update.
- Research & Academic writing skills.

13.3.3 Risk Management

The following table identifies possible risks the author could face and how they could mitigate them.

Table 5: Risk Management Plan

Risk Item	Severity	Frequency	Mitigation Plan
Lose access to development code	5	2	Backup code on source control and cloud storage.
Invalid hypothesis	3	2	Continue researching since the final output is a research contribution regardless.
Corrupted documentation	4	4	Store all necessary documentation on the cloud as well as external storage.
Inability to deliver all expected deliverables	4	2	Follow a list of priorities and deliver accordingly.
Lack of in-depth knowledge for ML algorithm development	6	3	Get insights from domain experts and, if necessary, the author of the proposed algorithm.

References

S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008. https://bitcoin.org/bitcoin.pdf.

Rahouti, Mohamed, et al. “Bitcoin Concepts, Threats, and Machine-Learning Security Solutions.” IEEE Access, vol. 6, 2018, pp. 67189–205. DOI.org (Crossref), https://doi.org/10.1109/ACCESS.2018.2874539.

Kervanci, I. sibel, and Fatih Akay. “Review on Bitcoin Price Prediction Using Machine Learning and Statistical Methods.” Sakarya University Journal of Computer and Information Sciences, Dec. 2020. DOI.org (Crossref), https://doi.org/10.35377/saucis.03.03.774276.

Answers to Your Forecasting Questions | Answers To Your Forecasting Questions 26 1201 | IBF. https://ibf.org/knowledge/jbf-articles/answers-to-your-forecasting-questions-26-1201.

Kenneth B. Kahn. How to measure the impact of a forecast error on an enterprise? The Journal of Business Forecasting Methods & Systems, 22(1), Spring 2003.

Makridakis, Spyros, et al. “Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward.” PLOS ONE, vol. 13, no. 3, Mar. 2018a, p. e0194889. PLoS Journals, https://doi.org/10.1371/journal.pone.0194889.

“The M4 Competition: Results, Findings, Conclusion and Way Forward.” International Journal of Forecasting, vol. 34, no. 4, Oct. 2018b, pp. 802–08. DOI.org (Crossref), https://doi.org/10.1016/j.ijforecast.2018.06.001.

Smyl, Slawek. “A Hybrid Method of Exponential Smoothing and Recurrent Neural Networks for Time Series Forecasting.” International Journal of Forecasting, vol. 36, no. 1, Jan. 2020, pp. 75–85. DOI.org (Crossref), https://doi.org/10.1016/j.ijforecast.2019.03.017.

Hasani, Ramin, et al. Liquid Time-Constant Networks. arXiv, 14 Dec. 2020. arXiv.org, https://doi.org/10.48550/arXiv.2006.04439.

Chen, Ricky T. Q., et al. Neural Ordinary Differential Equations. arXiv, 13 Dec. 2019. arXiv, https://doi.org/10.48550/arXiv.1806.07366.

Poulopoulos, Dimitris. “Is ‘Liquid’ ML the Answer to Autonomous Driving?” Medium, 1 Feb. 2021, https://towardsdatascience.com/is-liquid-ml-the-answer-to-autonomous-driving-bf2e899a9065.

Hochreiter, Sepp, and Jürgen Schmidhuber. “Long Short-Term Memory.” Neural Computation, vol. 9, no. 8, Nov. 1997, pp. 1735–80. DOI.org (Crossref), https://doi.org/10.1162/neco.1997.9.8.1735.

Oreshkin, Boris N., et al. N-BEATS: Neural Basis Expansion Analysis for Interpretable Time Series Forecasting. arXiv, 20 Feb. 2020. arXiv.org, http://arxiv.org/abs/1905.10437.

“Autoregressive Integrated Moving Average (ARIMA).” Investopedia, https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp.

Roy, Shaily, et al. “Bitcoin Price Forecasting Using Time Series Analysis.” 2018 21st International Conference of Computer and Information Technology (ICCIT), IEEE, 2018, pp. 1–5. DOI.org (Crossref), https://doi.org/10.1109/ICCITECHN.2018.8631923.

Rizwan, Muhammad, et al. “Bitcoin Price Prediction Using Deep Learning Algorithm.” 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), IEEE, 2019, pp. 1–7. DOI.org (Crossref), https://doi.org/10.1109/MACS48846.2019.9024772.

Fleischer, Jacques Phillipe, et al. “Time Series Analysis of Cryptocurrency Prices Using Long Short-Term Memory.” Algorithms, vol. 15, no. 7, July 2022, p. 230. DOI.org (Crossref), https://doi.org/10.3390/a15070230.

Saunders, M. N. K., et al. Research Methods for Business Students. 4th ed, Financial Times/Prentice Hall, 2007.

Falessi, Davide, et al. On the Need of Preserving Order of Data When Validating Within-Project Defect Classifiers. arXiv, 31 July 2020. arXiv.org, http://arxiv.org/abs/1809.01510.