Data-Driven Risk Forecasting and Algorithmic Trading Models for Cryptocurrencies

by

Japjeet Singh

A thesis submitted to
The Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements
of the degree of

Master of Science

Department of Computer Science
The University of Manitoba
Winnipeg, Manitoba, Canada
December 2022

Thesis advisor | Author

Dr. Ruppa K. Thulasiram (Advisor)

Dr. A. Thavaneswaran (Co-advisor) | **Japjeet Singh**

Data-Driven Risk Forecasting and Algorithmic Trading Models for Cryptocurrencies

Abstract

Capital investment is a key pillar of the modern economic system. It helps people to hedge their savings against inflation, while for businesses, it is an important means to raise capital. Various financial instruments like stocks, bonds, and derivatives have been invented to formalize such investments. These instruments are traded on financial exchanges worldwide and their prices are very dynamic. As a large volume of historical financial data becomes available, complemented by the growing availability of computing power, an increasing amount of literature within computational finance (CF) is focused on the development of investment management models using techniques from Statistics and Machine Learning. The aim of such models is to maximize profits while minimizing the associated risk.

Following the global financial crisis of 2008, regulatory authorities around the world mandated the implementation of stringent risk management controls for financial institutions. In the Basel III monitoring report by the Basel Committee on Banking Supervision published in 2021 [1], financial institutions are required to use more complex models for efficient capital allocation and enhanced risk management. Risk modelling is an active area of academic research within CF which involves developing models for quantifying and forecasting the uncertainty associated with financial market investments.

A new way of maintaining financial transfers by using a decentralized ledger, later called Blockchain, was introduced in the Bitcoin Whitepaper in 2008. The scope of blockchain-powered financial services has grown exponentially since then with the launch of different blockchain-powered digital assets, such as cryptocurrencies and digital tokens. However, unlike traditional financial assets, the risk modelling for these digital assets is rarely studied. My thesis research revolves around risk forecasting methods for blockchain-based digital assets in comparison with traditional assets like stocks. The volatility of returns is the key parameter used in all the widely used risk metrics.

The main contributions of this thesis lie in the application of novel data-driven forecasting models to forecast the risk of the price and algorithmic returns of cryptocurrency assets. In the studies carried out for my thesis, I have demonstrated the superiority of the data-driven volatility forecasting models over traditional models.

In the first study for this thesis, I proposed the data-driven and neuro-volatility risk forecasting models for Cryptocurrencies. The results are quantified with two widely used risk metrics, Value at Risk (VaR) and Expected Shortfall (ES). The data-driven and neuro-volatility forecasts demonstrated a significantly higher risk for cryptocurrencies than traditional technology stocks. Based on the real data experiments I observed that the data-driven models produced better forecasts for cryptocurrencies, while for the regular stocks and indexes, the neuro-volatility model gave better forecasts. Also, the data-driven models are more efficient in terms of computational complexity as the running time of the neuro-volatility model is significantly higher than that of the data-driven model. In an extension of this study on the intra-day minute frequency data, no definitive conclusion could be drawn due to the limited size of data available at this frequency.

In another study, I used a similar methodology to quantify and forecast the risk for the returns generated by an active algorithmic trading strategy (henceforth called "Algo returns") for stocks and cryptocurrencies. I used Sharpe Ratio (SR) as the metric to factor in the risk. The results of this study suggest that there is no significant difference in SR of algorithmic returns between stocks and cryptocurrencies. Based on the width of fuzzy intervals, I observed that the data-driven model generates better forecasts for cryptocurrencies, while for the regular stocks and indexes, no such definitive conclusion could be drawn.

In a separate study, I propose a novel active algorithmic trading strategy for cryptocurrencies (BTC and ETH), based on an LSTM price prediction model. The strategy consistently generates positive trading cashflows on backtesting, as validated by taking an average of multiple runs. I also implemented the strategy with an optimal ARIMA model for generating price forecasts for the strategy. The results, in terms of cumulative trading cash flow, demonstrate the superiority of LSTM forecasts over ARIMA forecasts. The strategy was also tested on a smaller dataset of minute frequency data and it gives a positive cash flow in that case as well. The fuzzy forecasts of Algo returns are significantly narrower when compared to a simple buy-and-hold strategy, signifying the stability in returns.

I also did a brief systematic review of the Decentralized Finance (DeFi) ecosystem, along with a case study comparing the market risk of holding digital tokens against providing liquidity to a Decentralized Exchange (DEX). The results show that liquidity provision to DEX reduces the market risk of investing in digital tokens.

The main findings from the studies indicate the superiority of the data-driven approach over the ones available in the literature. Despite being significantly efficient in terms of computation time, the data-driven approach gives equivalent or better risk forecasts than the neuro-volatility models in different studies.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to begin by thanking God, my advisor, committee, family, friends and everyone who have supported me along the way.

This thesis is dedicated to my Parents.

Chapter 1

Introduction and Motivation

Modelling and forecasting volatility is crucial for risk management in both traditional financial markets and cryptocurrency markets [3]. As per the Basel Committee report on Banking Supervision in 2021 [1], financial institutions are required to use more complex credit scoring models for efficient capital allocation and enhanced risk management. A large portion of the capital investment in today's economy is handled by a few very large financial institutions, which makes the system overly centralized. However, following the global financial crisis of 2008, an interesting development took place in the world of finance with the launch of Bitcoin payment network [4], which resulted in the launch of first *cryptocurrency*, Bitcoin. Cryptocurrency is a decentralized medium of exchange that relies on cryptographic primitives to facilitate the trustless transfer of value between different parties. Instead of being physical money, cryptocurrency payments exist purely as digital entries on an online ledger called Blockchain that describe specific transactions. Blockchain gained significant traction and many more such decentralized digital currencies followed, and the term *cryptocur-*

rency became prevalent. Cryptocurrencies derive their value purely from the trust that is placed on them, and they are not backed by any commodity, such as gold or silver. This makes the price of cryptocurrencies highly volatile and hard to forecast. Some advantages that cryptocurrencies offer over traditional payment methods (such as credit cards) include high liquidity, lower transaction costs, and anonymity[5]. Recently, there is a significant growth in research taking place in the area of risk forecasting for financial time series data. Three key metrics that characterize market risk are volatility, Value at Risk (VaR), and Expected Shortfall (ES) among many others. In financial econometrics, *volatility* is often defined as the measure of variation in the price of the series over time and mathematically stated by the standard deviation of logarithmic returns. *VaR* is a widely used risk metric, which represents the financial loss that can be incurred by an investment for a given level of confidence and over a certain time. *ES* is a risk metric that complements VaR, and is calculated by taking a conditional expectation of losses beyond the VaR cut-off.

1.1 Computing Fuzzy Risk Forecast Intervals

Different volatility forecasting models can be used to calculate VaR and ES. Some widely used models in practice include Historical Simulation (HS), Moving Average (MA), Exponentially Weighted Moving Average (EWMA), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) as well as student-t GARCH (tGarch). In these models the distribution of logarithmic returns on daily closing prices is generally assumed to be either Normal or Student's t with fixed degrees of freedom (d.f.) [6]. However, it has been shown in empirical studies that sometime the logarithmic returns may have heavy-tailed distributions, such as Student's t having less than four degrees of freedom, and theoretically infinite kurtosis [7]. In regular regression and volatility forecasting, more complicated models have been proposed to capture variations in the real world. However, a more complex and flexible model leads to the risk of over-fitting. This can be overcome by regularized estimates, which are widely used in different machine learning models to prevent over-fitting. These regularized methods are also known as shrinkage methods, in which data values are shrunk towards a central point, such as the mean. A recent study [8] explored the application of three different regularization methods i.e., Ridge, Lasso, and elastic net in financial risk forecasting on the stock prices of some large technology companies. They found that the regularized versions of data-driven models improved the stability of risk forecasts, as measured by the model risk ratio. Our study also explores the application of these regularized method for risk forecasting of cryptocurrencies. We use the concepts of fuzzy intervals to study the efficiency of different forecasting models. The use of fuzzy set theory to model certain financial problems [9], [10], [11] is of particular interest to several researchers due to its ability to quantitatively and qualitatively model the problems, which involve vagueness and imprecision. Recent studies have shown that a fuzzy random variable can be considered as a measurable mapping from a probability space to a set of fuzzy variables [12]. Fuzzy time series models provide a new avenue to deal with subjectivity observed in most financial time series models. Most of the fuzzy financial models developed so far have generally, been confined to modeling parameters through some form of defuzzification or linear type of fuzzy numbers such as Trapezoidal Fuzzy Number (Tr.F.N.) or Triangular Fuzzy

Number (T.F.N.). The main reason for using a linear membership function is to avoid complex nonlinear computations [13], [14]. Fuzzy methods remain difficult to use in practice, and hence, there is a need for data-driven approaches to pragmatically use the fuzzy models for real-world financial models. In this study, we use the fuzzy approach in conjunction with the data-driven volatility forecasting models. The data for this study includes the price data of five stocks/indexes and six cryptocurrencies. The first set includes Apple (APPL), Amazon (AMZN), Facebook (FB), Google (GOOG), and CBOE Volatility Index (VIX), while the second set includes Bitcoin (BTC), Ether (ETH), Binance Coin (BNB), Rippe (XRP), Dogecoin (DOGE), and Cardano (ADA) cryptocurrencies. The selection of cryptocurrencies is based on their market capital as per Coinmarketcap. Even though Tether, Solana, and Polkadot are among the top six cryptocurrencies by market capitalization, they are not included in the study. Tether (USDT) is a stable coin and exhibits negligible volatility, while Solana (SOL) and Polkadot (DOT) were recently launched, and their price data is unavailable for the period of study. We collect all our data from the Yahoo! Finance, which is a leading source of financial data. Downloadable data include opening, high, low, closing, and adjusted prices as well as volume for numerous stocks, indexes, and cryptocurrencies. In this study, we specifically use the daily adjusted price, which is an amended version of the closing price obtained by nullifying the impact of certain corporate actions which can affect the price post market closing. The window of our study period is from 2017-10-01 to 2021-11-26, to accommodate most of the cryptocurrencies with large market capitalization, as many of them are fairly new.

1.2 Data-Driven Fuzzy Forecasts for Cryptocurrencies

Due to strict risk management requirements, financial institutions at large do not invest in cryptocurrencies, as existing risk management modelling is not sufficient to capture the uncertainty associated with cryptocurrency volatility forecasts. Fuzzy confidence intervals [15] can prove effective in incorporating inherent uncertainty in the forecast of cryptocurrency returns.

For any financial asset, the most common way to quantify return is to capture the day-to-day changes in its *adjusted closing price*. The adjusted closing price is an asset's closing price amended to reflect the stock's value after accounting for any corporate action(s). The return metric used in this work is called *Algo Return*, which is calculated by algorithmically simulating long and short positions at different points of time. The Simple Moving Average (SMA) of the asset price, which is an average of the past adjusted closing prices P_t, $t = 1, 2, ..., D$ is a key parameter to simulate the trading strategy. The short-term and long-term SMA trend indicators are calculated to decide whether to buy or sell an asset at a given time point. The trend-following position indicator takes the values of 1 and -1 at each trending time t. If the short-term SMA $\geq$ long-term SMA, then the position is selected as long position (1), and if the short-term SMA $<$ long-term SMA, then the position is selected as short position (-1).

The daily adjusted price can be converted to simple returns as $R_t = (P_t - P_{t-1})/P_{t-1}$. When the return is multiplied by the corresponding position for each t, the resulting quantity is called Algo return (A_t). If μ_A and σ_A are the mean and standard deviation of daily Algo returns respectively, then the daily Sharpe ratio (SR) is computed as

$$Daily.SR = \frac{E(A_t - r_f/N)}{\sqrt{Var(A_t - r_f/N)}} = \frac{\mu_A - r_f/N}{\sigma_A} \tag{1.1}$$

where r_f is the annual risk-free rate, and N is the number of trading periods in a year. Sharpe ratio helps investors to understand the return of an investment compared to its risk. A higher SR is indicative of higher risk-adjusted returns for an investment, and hence it can also be used as a measure of the quality of an algorithmic trading strategy. A strategy that gives high returns and low volatility is always preferred.

Two different study periods are considered for stocks/indexes and cryptocurrencies due to different daily trading periods for a given year. For stocks/indexes, weekend adjusted closing prices are not available. However, for cryptocurrencies, weekend-adjusted closing prices are available. Thus, for a given year, cryptocurrencies have 365 daily trading periods ($N = 365$), and stocks/indexes have only 252 daily trading periods ($N = 252$). Therefore, to accommodate most of the cryptocurrencies with large market capitalization and most recent data, the study periods of 2017-01-01 to 2021-12-31 and from 2018-01-01 to 2021-12-31 have been chosen for stocks/indexes and cryptocurrencies, respectively.

1.3 Novel Algorithmic Trading Strategy for Bitcoin and Ethereum

During the past decade, Deep Neural Network (DNN) based models have gained popularity in diverse machine learning problems, such as speech recognition, natural language processing, computer vision, robotics, computational finance, etc. Moreover, Neural network based sequence models like LSTM [16], GRU [17] and Transformers [18] have the capability to model the prices of financial assets. This capability has been widely explored in recent years with the general rise of deep learning in traditional finance, while its use in a decentralized currency system has been limited. Also, there is a growing interest in using neuro volatility models for volatility forecasting of financial time series [19].

Though there has been high volatility in the cryptocurrency market and recent slumping price in Bitcoin, experts suggest that Bitcoin could cross the $100,000 mark, soon possibly in a year or two. A recent study by Deutsche Bank found that about a quarter of Bitcoin investors believe Bitcoin prices will be over $110,000 in five years (See - [20]). Since such public survey based predictions are unscientific, there are efforts to apply ML techniques used for traditional financial asset price prediction to cryptocurrencies as well.

This part of the thesis work explores the use of LSTM based price prediction model for formulating a trading strategy for Bitcoin, the most popular and valuable cryptocurrency. The regression-based one-step ahead stock price prediction primarily uses the historical closing prices as input variables to predict the future price in the next step. In addition, the techniques from the fuzzy set theory are used to quantify the stability of the proposed algorithm. In particular, the fuzzy interval widths using alpha-cuts of a non-linear adaptive fuzzy membership function are provided for the algorithmic returns generated using the proposed strategy and a comparative analysis with a simple buy-and-hold strategy.

The rest of the thesis is structured as follows: In Chapter 2, the related work for the data-driven models for forecasting financial market risk for both traditional financial stocks and cryptocurrencies are summarized. Chapter 3 provides the theory of financial risk modelling and forecasting. The methodology and implementation of experiments performed for the thesis are presented in Chapter 4. Chapter 5 summarize the results of these experiments. Chapter 6 is an extension of this work to study the market risk in a recently emerging field of digital tokens and Decentralized Finance (DeFi) markets. We draw our conclusions from these studies and summarize them in chapter 7.

Chapter 2

Related Work

The relevant academic literature, related to each of the key problems described in Chapter 1 along with its contributions are presented in this chapter.

2.1 Traditonal Data-Driven Approaches of Risk Forecasting

Neuro-fuzzy volatility forecasting models can prove highly instrumental in the area of computational finance as they can be applied to problems across risk management, hedging and pricing of financial instruments. The review by Poon and Granger [21] studies various volatility forecasting models, defining volatility as a sample standard deviation of logarithmic returns. The most widely used models for volatility forecasting are the ARCH class of models introduced by Engle in his seminal work [22] and reviewed by Bera and Higgins [23]. A commonly used metric to quantify market risk is Value at Risk (VaR), which is the expected value of loss incurred in a given time interval and given confidence level [24]. Recently, some interesting work has been reported by Thavaneswarana et al. ([7], [25]) in the area of data-driven statistical models for financial time series. One such method is the novel data-driven generalized Exponentially Weighted Moving Average (DD-EWMA) model to estimate VaR forecasts for returns of stocks and indexes with larger kurtosis. Another promising area is integrating the core concepts of neuro-fuzzy prediction with the volatility forecasting models to get better estimates [25]. This approach has been demonstrated in [26] for applications in option pricing, and in [27] for applications in portfolio optimization. This data-driven model was further improved by Liang et al.[8] by introducing regularization to estimate the optimal model parameters. A large portion of the literature on volatility forecasting targets the returns from stocks or indexes. However, an entirely new financial asset called cryptocurrency has emerged with introduction of Bitcoin - a decentralized digital currency [4]. Some key features of cryptocurrencies can be found in a technical survey by Tschorsch and Scheuermann [28]. Caporale and Zekokh used the Markov-switching GARCH models to model the volatility of the four most popular Cryptocurrencies at the time of its writing [29]. Recently, a significant stream of research literature has focused on using machine learning models for predicting the future price movements in cryptocurrencies using the historical prices and other extraneous factors as inputs to models (see [30], [31]). Liu and Tsyvinski [32] review different theoretical models used to describe cryptocurrency prices and lay down a set of tenets for benchmarking such models in the future.

Several previous works [33], [34] found hybrid of GARCH and neural network based model to outperform the vanilla GARCH models in diverse set of assets, such as gold and NASDAQ indices.

A recent paper by Rahimikia and Poon [35] describes a Long Short term Memory (LSTM) neural network based volatility forecasting model for some popular NASDAQ stocks. The model has been primarily trained on features extracted from the LOBSTER limit order book dataset [36]. In addition, features extracted from the news articles of the major financial media outlets were incorporated to augment the model performance.

2.2 LSTM Based Forecasting

Long short-term memory (LSTM) is one of many types of recurrent neural network (RNN), which uses data from past and use them to predict future data [37]. In general, an Artificial Neural Network (ANN) consists of three layers: input layer, Hidden layers, output layer. There are many research efforts to predict future value or movement of stock prices using technical analysis of historical data. Most of the research formulate the problem of the future prediction in either of the two categories, future movement prediction or future value prediction. First one is formulated as classification problem and the other as regression problem. The published literature in this area can further be broadly split into three types of approaches: multivariate regression (see e.g., [38]), time series based prediction like Granger causality (see e.g., [39]) and machine learning based algorithms(e.g., [40]). Performance of ML models have, in general, been found to be superior on financial time series data in comparison to the models of the first two categories. However, when significant randomness is exhibited in time series data the forecast accuracies of these models would decrease significantly. *Cryptocurrencies* are a novel financial asset class based on the Blockchain decentralized ledger technology. *Bitcoin* was the first true cryptocurrency, and it was introduced in a white paper [4] published under the pseudonym of Satoshi Nakamoto. Some academic works [41], [42] consider cryptocurrencies as a new asset class instead of currency due to their highly volatile prices when compared to other fiat currencies. The academic literature pertaining to cryptocurrencies has also grown rapidly with their rising market capitalization and investor interest in cryptocurrencies. The Deutsche Bank research report [43] published in year 2020 analyses the increasing mainstream adoption of cryptocurrencies and other blockchain based digital assets. Many studies in recent years have tried to analyze the economic and properties of popular cryptocurrencies, such as volatility in market price [44]; dynamic correlation between different cryptocurrencies [45]. Vidal-Tomás and Ibañez [46] carried out a study to determine efficiency in the Bitcoin market to conclude that the markets are semi-efficient and over time the efficiency seems to be increasing.; Urquhart[47] studied the Bitcoin prices to analyze price clustering and found significant statistical evidence supporting clustering at round numbers. The price forecasting of cryptocurrencies, particularly Bitcoin, is an area that has allured significant academic interest in the last few years.There are two categories of forecasting models which have been explored in this area - the traditional statistical models like *ARIMA* and the more recently developed neural network based models like *Recurrent Neural Network (RNN)*. Rebane et al. [48] compare the performance of ARIMA with *seq2seq*, a variant of recurrent neural network in task of forecasting Bitcoin prices across different time intervals, and demonstrated a superior performance of seq2seq.

Wu et al. [49] further established the efficacy of an LSTM model for daily price forecast of Bitcoin by statistically proving their superiority over traditional models. A more sophisticated use of neural network models was proposed in [50], with the use of technical indicators as input variable to *CLSTM*, a hybrid of Convolutional Neural Networks (CNN) and LSTM. Another interesting category of literature on cryptocurrency price forecasting explores incorporating extraneous data sources like Google trends [51] or Twitter feeds [52] as a proxy for investor sentiment.

The literature associated with the emerging area of decentralized finance (DeFi) is reviewed in Chapter 6 along with other studies related to DeFi.

Chapter 3

Financial Risk Modelling and Forecasting

Traditional and data-driven approach for risk forecasting are presented in this chapter. Also presented is LSTM based approach for Algorithmic trading in Cryptocurrencies.

3.1 Value-at-Risk (VaR) and Expected Shortfall (ES)

Throughout the finance literature, risk forecasts are calculated using a time series of logarithmic returns $r_t = \log P_t - \log P_{t-1}$ as an input to the forecasting model. In this work, the one-step-ahead forecasts of VaR and ES are computed using given equations (3.1) and (3.2):

$$\mathrm{VaR}_{t+1}(p) = -\hat{\sigma}_{t+1} F_r^{-1}(p), \tag{3.1}$$

$$\mathrm{ES}_{t+1}(p) = -\frac{\hat{\sigma}_{t+1}}{p} \int_{\infty}^{F_r^{-1}(p)} x f(x) dx. \tag{3.2}$$

Here $f(x)$ is the density function for the conditional distribution of logarithmic returns r_t, and $F_r^{-1}(p)$ is the inverse of Cumulative Density Function (CDF) of r_t evaluated at the tail probability p, while $\hat{\sigma}_{t+1}$ represent the volatility forecast at time $t+1$ given the past t observations. Thus, the process for forecasting the VaR and ES involves forecasting the volatility, and identifying a suitable distribution for log returns.

3.2 Data-Driven Risk Forecasting

3.2.1 Data-Driven Volatility Forecasting

As shown in [8] and [10], if a random variable follows a Student's t distribution with d.f. greater than two, we can compute the optimal value of d.f. ν by finding solutions of the following equation:

$$2\sqrt{\nu - 2} = (\nu - 1)\rho_X \mathrm{Beta}\left[\frac{\nu}{2}, \frac{1}{2}\right]. \tag{3.3}$$

Here ρ_X denotes the sign correlation of a random variable with its mean μ, and can be computed using below equation:

$$\begin{aligned} \rho_X &= \mathrm{Corr}(X - \mu, \mathrm{sgn}(X - \mu)) \\ &= \frac{E|X - \mu|}{2\sigma\sqrt{F(\mu)(1 - F(\mu))}} \end{aligned} \tag{3.4}$$

where $F(\mu)$ is the Cumulative Density Function (CDF) of X evaluated at the mean, and σ^2 is the finite variance of X. The method to find the conditional distribution of log returns for VaR forecasts has been discussed in [7]. Furthermore, the d.f. of student's t or generalized t distribution can be computed by using the sample estimate of sign correlation in equation (3.13). The observed volatility based on data-driven sign correlation incorporates skewness and non-normality.

3.2.2 Regularized Risk Forecasting

A further improvement to the data-driven volatility forecasts described above was suggested in [8] by introducing regularization to the volatility estimation model. After calculating the volatility forecasts $\hat{\sigma}_{t+1}$ using the data-driven approach, these forecasts are regularized with penalties using the elastic net regularization technique (3.5). The explicit form of the regularized volatility forecasts, $\tilde{\sigma}^{en}_{t+1}$, based on the elastic net penalties is given by:

$$\tilde{\sigma}^{en}_{t+1} = \frac{\mathrm{sgn}(\hat{\sigma}_{t+1} - s)(\mid \hat{\sigma}_{t+1} - s \mid -\omega^{en}\lambda)_+}{1 + (1 - \omega^{en})\lambda} + s, \tag{3.5}$$

where $\lambda, \omega^{en} \in [0, 1]$ are the tuning parameters of the model, and s represents the sample standard deviation of the time series.

3.2.3 Neuro Volatility Forecasting

A neural network (NN) is a powerful prediction model which can approximate any nonlinear real function on a bounded domain to a very high accuracy. The motivation for the design of neural networks comes from the interconnected neurons in the animal brain. A neural network fundamentally consists of the layers of nodes called *perceptrons*, and each node functions as a binary classifier. The simplest form of NN is a feed-forward neural network. It contains an input unit that reads the input variables, followed by an arbitrary number of interconnected hidden layers followed by an output layer. The transformations between two consecutive layers can be represented by the equations 3.6 and 3.7.

$$z_l^k = w_{l0}^{(k-1)} + \sum_{j=1}^{p_k-1} w_{lj}^{(k-1)} a_j^{(k-1)} \tag{3.6}$$

$$a_l^{(k)} = g^{(k)}\left(z_l^{(k)}\right) \tag{3.7}$$

Here, z_l^k is the linear transformation for l^{th} unit of the k^{th} layer of a feed forward NN, $w_{l0}^{(k-1)}$ is the bias term in $(k-1)^{th}$ layer, while $w_{lj}^{(k-1)}$ represent the weights of $(k-1)_{th}$ layer. $g^{(k)}$ is any non-linear function (eg: Sigmoid function). NNs differ from traditional time series forecasting models used in finance by the number of parameters that be tuned, which is much higher in NNs. Also, unlike traditional models, all the parameters do not need to be optimized in a NN to get a universal approximate solution. Most of the NN models studied in the finance literature deal with stock price prediction. Thavaneswaran et al.[25] were the first to study the volatility and VaR forecasting using a generalized neuro volatility model. They trained their neuro volatility model, a feed forward neural network, on the p lagged values of the centered absolute log returns $|r_{t-1}^*|$, $|r_{t-2}^*|$, $|r_{t-3}^*|$, ... , $|r_{t-p}^*|$ to predict the target variable $|r_t^*|$, which is defined as $r_t^* = \frac{r_t - \bar{r}}{\rho}$, as a prediction from the output layer of neural network. In our neuro volatility model, we use the p-lagged values of the volatility of absolute log returns V_{t-1}, V_{t-2}, V_{t-3}, ... , V_{t-p} as inputs and predict the target variable V_t as an output of the model (Fig. 3.1). To fit this neuro volatility model, we use the

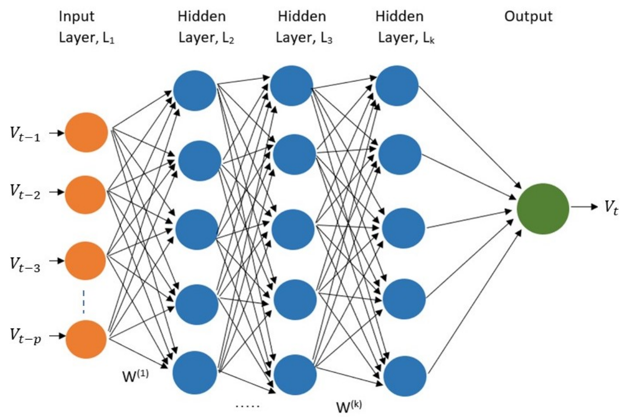

Figure 3.1: Illustration of feed-forward neuro volatility network

`nnetar` function of the R Package *forecast* [53].

3.2.4 Data-Driven Fuzzy Volatility Using Nonlinear Adaptive Fuzzy Numbers

If R be the set of all real numbers, a fuzzy number $A(x)$, $x \in R$ is of the following form

$$A(x) = \begin{cases} g(x) & when & [a, b) \\ 1 & when & [b, c) \\ h(x) & when & [c, d) \\ 0 & otherwise & \end{cases} \tag{3.8}$$

where g is a real, right continuous, and increasing function while h is a real, left continuous, and decreasing function. Also, a, b, c, d are real numbers and $a < b < c < d$. It should be noticed that $A(x)$ is a fuzzy number having strictly monotone shape functions as mentioned by Bodjanova [11]. (Readers are referred to Dubois and

Prade [54] and Zimmermann [55] for more details on fuzzy numbers. A fuzzy number A with shape functions g and h is denoted by $A = [a, b, c, d]_{m,n}$. Here, g and h are defined as:

$$g(x) = \left(\frac{x-a}{b-a}\right)^m \tag{3.9}$$

$$h(x) = \left(\frac{d-x}{d-c}\right)^n \tag{3.10}$$

Here, m and n are parameters of shape functions. In a special case where m and n equal 1, $A(x)$ can be simply written as $[a, b, c, d]$, a trapezoidal fuzzy number, which is a *Linear Fuzzy number*. When either m or n does not equal 1, the resulting fuzzy number $[a, b, c, d]_{m,n}$ is a modified form of trapezoidal fuzzy number known as nonlinear adaptive asymmetric fuzzy number (Figure 3.2).

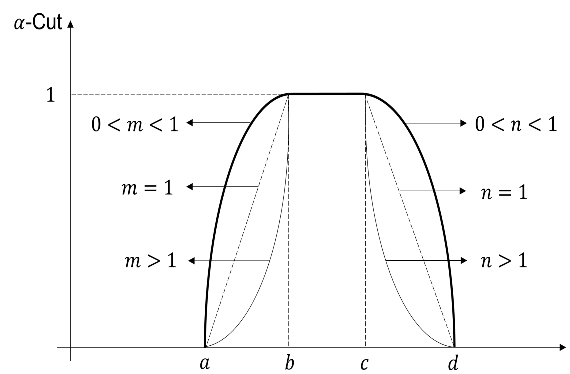

Figure 3.2: Nonlinear membership function used in this study which is a modification of trapezoidal membership function

To find the forecast interval widths of VaR and ES for stocks/indexes and cryptocurrencies, the α-cuts[1] of the annualized volatility is calculated using different α, (a, b, c, d), m, and n values. For this study we choose (a, b, c, d) as the 0.05, 0.25, 0.75,

[1]For a fuzzy number $A(x)$, its $\alpha - cuts$ are defined as, $A_\alpha = [a(\alpha), b(\alpha)]$, $a(\alpha)$, $b(\alpha) \in R$, $\alpha \in [0, 1]$ and $A_\alpha = [g^{-1}(\alpha), h^{-1}(\alpha)]$, $A_1 = [b, c]$, $A_0 = [a, d]$. If, $A = [a, b, c, d]_{m,n}$ then, $\forall \alpha \in [0, 1]$, $A_\alpha = [a + \alpha^{\frac{1}{m}}(b - a), d - \alpha^{\frac{1}{n}}(d - c)]$.

and 0.95 quantiles respectively of the volatility forecasts.

3.3 Fuzzy Risk Forecasting of Algo returns

The techniques from the Fuzzy set theory can be used to quantify the stability of risk forecast of an algorithmic trading strategy. In particular, the fuzzy interval widths calculated using the alpha-cuts of a non-linear adaptive fuzzy membership function on the algorithmic returns can be used as a metric for stability of forecast, and thus a proxy for market risk of the proposed algorithm. This method can be used to do a comparative analysis of different algorithmic trading strategies on a common asset while backtesting the returns.

3.3.1 Volatility Forecasts

The forecast of Sharpe Ratio (SR)[2] is carried out using the DD-EWMA using the *sign correlation* [10].

The sign correlation of a random variable X with mean μ is defined as,

$$\rho_X = \mathrm{Corr}(X - \mu, \mathrm{sign}(X - \mu)) \tag{3.12}$$

If X follows a Student's t distribution with sign correlation ρ_X and finite variance, the corresponding degrees of freedom (d.f.) ν can be computed by solving the

[2]Sharpe ratio is one of the most commonly used measures of *risk-adjusted returns for an investment strategy or portfolio. If μ_A and σ_A are the mean and standard deviation of returns respectively, then the Sharpe ratio (SR) is computed as*

$$SR = \frac{E(A_t - r_f/N)}{\sqrt{Var(A_t - r_f/N)}} = \frac{\mu_A - r_f/N}{\sigma_A} \tag{3.11}$$

following equation:

$$2\sqrt{\nu - 2} = (\nu - 1)\rho_X \text{Beta}\left[\frac{\nu}{2}, \frac{1}{2}\right] \tag{3.13}$$

The data-driven algorithmic volatility estimator, in terms of algorithmic returns $A_1, \cdots, A_n$, is given as,

$$\hat{\sigma}_A = \frac{1}{n}\sum_{t=1}^{n} \frac{|A_t - \bar{A}|}{\hat{\rho}_A} \tag{3.14}$$

where $\hat{\rho}_A$ is the sample sign correlation of A_t which can be calculated using 3.12. The asymptotic variance of the data-driven algorithmic volatility estimator $\hat{\sigma}_A$ is,

$$\left(\frac{1 - \rho_A^2}{\rho_A^2}\right)\frac{\sigma_A^2}{n} \tag{3.15}$$

The most commonly used volatility estimator in practice is the sample standard deviation s_n, and its asymptotic variance is given by,

$$\frac{(\kappa + 2)}{4}\frac{\sigma_A^2}{n}$$

where κ is the excess kurtosis. As reported by Thavaneswaran et al. [10]), investment returns follow Student's t distribution with degrees of freedom, d.f., less than four, and hence have an infinite kurtosis as per theory. Thus, the sample standard deviation estimator has a very large or infinite asymptotic variance for the returns with heavy-tailed distributions, such as returns from the technology stocks. For such distributions, Thavaneswaran et al. [10] propose an alternative data-driven volatility estimator $\hat{\sigma}_A$, which has a smaller and finite asymptotic variance.

Volatility forecasting has been widely used in risk management and asset pricing, but its use in algorithmic trading is not common. However, very recently, Thavaneswaran et al. [10] have discussed the usage of SR fuzzy forecasts based on volatility forecasting models.

For algorithmic return A_t, the conditional mean and conditional variance is defined as

$$E(A_t|\mathcal{F}_{t-1}) = \mu_t, Var(A_t|\mathcal{F}_{t-1}) = \sigma_t^2, t = 1, \cdots, n$$

where $\mathcal{F}_{t-1}$ is the data until time $t-1$. Further, as per [10], the volatility forecasting model for algorithmic returns can be written as,

$$\hat{\sigma}_{t+1} = (1-\omega)\,\hat{\sigma}_t + \omega\frac{|A_t - \bar{A}|}{\hat{\rho}_A}, \quad 0 < \omega < 1 \tag{3.16}$$

3.3.2 Annualized Sharpe Ratio (ASR) Forecasts

The data-driven estimates for SR of the algorithmic returns are computed using the SMA crossover strategy as described in [10]. We further compute the annualized SR (equation 3.17) from daily SR as it provides a better metric for comparing different assets.

$$Annualized.SR = \sqrt{N} Daily.SR \tag{3.17}$$

The α-cuts for σ_A can be written as,

$$\overline{\sigma_A(\alpha)} = \hat{\sigma}_A \pm cv_\alpha \sqrt{\frac{(1-\hat{\rho}_A^2)\hat{\sigma}_A^2}{n\hat{\rho}_A^2}} \tag{3.18}$$

where cv_α are the critical value of level α and $\hat{\sigma}_A$ is the data-driven volatility estimate (DDVE) of algorithmic returns. Then, the α-cuts of annualized SR are written as,

$$\left(\frac{\sqrt{N}(\bar{A}_t - r_f/N)}{UL^{\sigma_A}}, \frac{\sqrt{N}(\bar{A}_t - r_f/N)}{LL^{\sigma_A}}\right) \tag{3.19}$$

where LL^{σ_A} denotes lower limit of σ_A and UL^{σ_A} denotes upper limit of σ_A.

3.3.3 Dynamic Data-Driven Rolling Sharpe Ratio Forecasts

Time-varying volatility models would be more appropriate if there is an indication of significant sample autocorrelation in the absolute values of the algorithmic returns. The first step toward designing such a model is to compute the daily DD-EWMA volatility forecasts and daily data-driven neuro volatility forecasts.

The sample sign correlation $\hat{\rho}_A$ and observed algorithmic volatility $Z_t = |A_t - \bar{A}|/\hat{\rho}_A$ can be computed based on the past observations of algorithmic returns. The optimal smoothing constant ω is determined by minimizing the one-step ahead forecast error sum of squares (FESS). Using the optimal ω, smoothed value S_t ($S_t = \omega Z_t + (1-\omega)S_{t-1},\ t = 1, \ldots, k$) can be calculated recursively.

The initial smoothed value is calculated using the first l observations, and the last smoothed value of S_t is the one-day-ahead volatility forecast. The root mean square error (RMSE) of volatility forecasts can be computed as,

$$\sum_{t=l+1}^{k} (Z_t - S_{t-1})^2/(k-l)$$

We use the rolling window approach to calculate daily and annualized SR forecasts using the algorithmic DD-EWMA volatility method and neuro volatility forecasts. The neuro volatility forecasts in our thesis have been computed by using the *nnetar* function from R package *forecast* [53]. Also, the α-cuts of the SR forecasts are computed using the below equation.

$$\begin{aligned}\overline{Daily.SR(\alpha)} &= (LL^{Daily.SR}, UL^{Daily.SR}) \\ &= mean(Daily.SR_i) \pm cv_\alpha sd(Daily.SR_i)\end{aligned}$$

The α-cuts of annualized SR forecasts can be calculated by daily SR forecasts as,

$$\overline{Annualized.SR(\alpha)} = (\sqrt{N}\, LL^{Daily.SR}, \sqrt{N}\, UL^{Daily.SR})$$

Chapter 4

Methodology and Implementation for Computing Risk Forecasts

The proposed algorithms, experimental methodology, and details of implementation are discussed in this chapter.

4.1 Fuzzy Risk Forecasting of Asset Returns

This study was carried out to apply the data-driven and neuro-fuzzy forecasting models to forecast the market risk of Stocks and Cryptocurrencies using Value at Risk (VaR) and Expected Shortfall (ES) as metrics for risk. In this study, we chose a period from October 1 2017 to November 26 2021, even though we use only the last 1000 days of adjusted closing prices to calculate log returns for both stocks and cryptocurrencies. It is important to note that the adjusted closing prices are available throughout the week for cryptocurrencies, however, for stocks, adjusted closing prices are only available for weekdays.

For the more recent extension of this study to high-frequency intra-day price fluctuations, we use the data for September 29, 2022. The ticker symbol for FB has been changed to META which is reflected in this extended study.

Also, while running the neuro-volatility model, 63-day rolling forecasts and 90-day rolling forecasts were considered to cover three months period for stocks and cryptocurrencies. We also designed an algorithm for risk forecasting (Algorithm 1) in this study.

4.2 Risk Forecasting of Algo Returns

For this study, we consider a simple yet effective algorithmic trading strategy called Simple Moving Average (SMA) crossover strategy. This strategy calculates two separate simple moving averages (SMAs) with two different window sizes. The crossing of these two SMAs signals a change in strategy (buy/sell). Based on the strategy, one would sell an asset if the SMA with shorter window size crosses below the SMA with longer window size (green shaded area in Figure 4.1). Contrarily, one would buy an asset if SMA with shorter window size crosses above the SMA with longer window size (pink shaded area in Figure 4.1). Considering adjusted closing prices of AAPL, the SMA crossover strategy with a short-term window size of 10 days and long-term window size of 20 days is visualized in Figure 4.1.

In order to select the optimal SMA window sizes for different assets, the annualized SR estimates are used. We use the *brute force* technique to find the optimal short-term window size by fixing the long-term window size to values in 20, 30, 60.

Algorithm 1 Risk forecasts

Require: Data: adjusted closing price of stock/index and Cryptocurrencies $P_t, t = 0, \ldots, n$

1: $r_t \leftarrow \log P_t - \log P_{t-1}, t = 1, \ldots, n$

2: $\hat{\rho} \leftarrow \mathrm{Corr}(r_t - \bar{r}, \mathrm{sgn}(r_t - \bar{r}))$ {Compute sample sign correlation of r_t}

3: $\nu \leftarrow$ Solve $2\sqrt{\nu - 2} = \hat{\rho}(\nu - 1)\mathrm{Beta}\left[\frac{\nu}{2}, \frac{1}{2}\right]$ {Determine t d.f. of the conditional distribution of r_t}

4: $V_t \leftarrow |r_t - \bar{r}|/\hat{\rho}$ {Compute forecasts of volatility}

5: $s \leftarrow \mathit{mean}\,(|r_t - \bar{r}|/\hat{\rho})$ {Compute the sample mean of forecasts of volatility}

6: $S_0 \leftarrow \bar{V}_k$ {Initial volatility forecast}

7: $\alpha \leftarrow (0, 1)$ {Set a range for the smoothing constant}

8: $S_t \leftarrow \alpha V_t + (1 - \alpha)S_{t-1}, t = 1, \ldots, n$ {Calculate smoothed value of volatility}

9: $\alpha_{opt} \leftarrow \arg\min_\alpha \sum_{t=k+1}^{n}(V_t - S_{t-1})^2$ {Determine the optimal value of α}

10: $S_t^{dd} \leftarrow S_t, t = 1, \ldots, n$ {Get Smoothed value}

11: $(\hat{\sigma}_{t+1})_{dd} \leftarrow S_n^{dd}$ {Get the data-driven estimate of volatility}

12: $(\hat{\sigma}_{t+1})_{en} \leftarrow S_n^{en}$ {Get the elastic-net regularized data-driven volatility estimate (refer [8]}

13: Pass $V_1 \ldots, V_1$ to the NN using `nnetar` function and assign the output of NN, i.e., $\hat{V}_{t+1}$ to $\hat{\sigma}_{t+1}$ {Get the neuro-volatility estimate}

14: Calculate the α-cuts of the annualized volatility for the data-driven and neuro-volatility forecasts {Use the non-linear fuzzy number defined in 3.2.4 }

15: $\hat{\mathrm{VaR}}_{n+1} \leftarrow -\hat{\sigma}_{n+1}F_r^{-1}(p)$ {VaR forecast from volatility}

16: $\hat{\mathrm{ES}}_{n+1} \leftarrow -\frac{\hat{\sigma}_{n+1}}{p}\int_{-\infty}^{F_r^{-1}(p)} xf(x)dx$ {ES forecast from volatility}

17: **return** $\hat{\mathrm{VaR}}_{n+1}, \hat{\mathrm{ES}}_{n+1}$

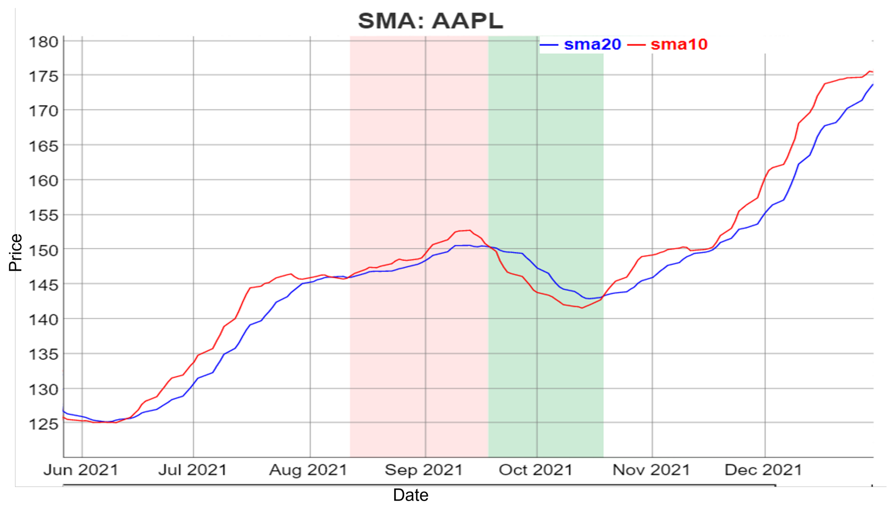

Figure 4.1: SMA crossover strategy executed on the AAPL stock for period of study. The start of green shaded area signals sell while the pink shaded area signals buy

The corresponding short-term window is selected from a range of 1 to the long-term window size. Figure 4.2 shows the progression of SR for AAPL. It can be seen that the SR stays relatively stable after the short-term window size of 10, and after the window size of 20, SR becomes almost constant, and hence it is chosen as the long-term window size. Similarly, for BTC, if we choose a long-term window size of 30, then 10 can be selected as a short-term window size.

As observed empirically, for all the stocks/indexes, the short-term window size can be chosen as 10 with the long-term window size of 20. For cryptocurrencies, short-term window sizes vary between 10 and 20, while 30 is an appropriate long-term window size. Our observation suggests short-term window sizes of 10 for ETH, 15 for BNB and XRP, and 20 for DOGE and ADA are appropriate.

Considering the adjusted closing prices of the assets, we proposed an algorithm (Algorithm 2) to compute the respective daily algorithmic returns (A_t). For all the stocks/indexes, the short-term and the long-term window sizes are chosen to be 10 and

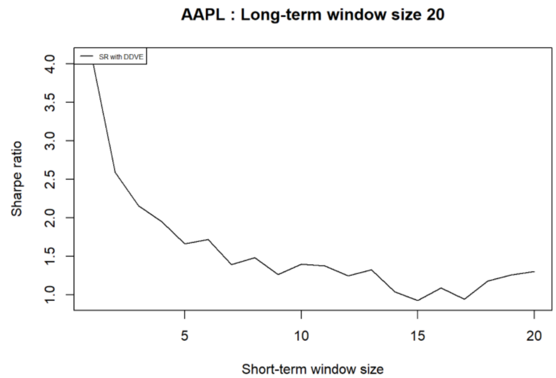

Figure 4.2: Estimated Annualized SR for AAPL plotted against the short-term window size to choose the best possible window size

Algorithm 2 SMA crossover trading strategy

Require: $P_t, t = 0, \ldots, n$ (Adjusted closing price of the asset), S (short-term window size), L (long-term window size)

$R_t \leftarrow (P_t - P_{t-1})/P_{t-1}, t = 1, \cdots, n$ {Calculate returns}

$SSMA_t \leftarrow SMA(P_t, S)$ {Short-term SMA}

$LSMA_t \leftarrow SMA(P_t, L)$ {Long-term SMA}

if $SSMA_t \geq LSMA_t$ **then**

  $Position_t \leftarrow 1$

else

  $Position_t \leftarrow -1$

end if

$AlgoReturn \leftarrow R_t * Position_t$ for each t {Algorithmic returns}

return $AlgoReturn$

20, respectively. Even though different short-term window sizes seem appropriate for different cryptocurrencies, for simplicity, the short-term window size of 20 is chosen for all the cryptocurrencies. Nonetheless, the long-term window size of 30 remains fixed for all the cryptocurrencies.

We then compute the summary statistics of the algorithmic return A_t for all the assets being studied. This is followed by the calculation of the annualized SR (ASR) and daily algorithmic volatility estimates based on the asymptotic variance of the data-driven volatility estimator. Similarly, sample standard deviation (s_n), mean absolute deviation (MAD) ($\hat{\rho}_A s_n$), and Value at Risk at $\alpha = 0.05$ ($VaR_{0.05}$) based on t distribution of A_t are used to estimate the daily volatility and annualized SR.

We obtain the rolling forecasts of the daily and annualized SR. The most recent 1000 data points of algorithmic returns are chosen to calculate the rolling forecasts of volatility. We compute the one-day ahead algorithmic volatility forecast and RMSE using the window sizes of 60 and 90 for stocks/indexes and cryptocurrencies, respectively. The selected window sizes account for three months of consecutive observations. For each asset, we define the *daily SR forecast* as the average of one-day-ahead SR forecasts computed for each of the rolling windows. we also compute the daily and annualized rolling SR forecasts using the neuro volatility model, for which we use similar window sizes as the data-driven rolling forecast.

4.3 Proposed Novel Algorithmic Trading Strategy for Cryptocurrencies Based on Deep-Sequential Model

In this study, we propose and implement a novel algorithmic trading strategy based on price predictions of a Long Short term Memory based forecasting model.

We also provide the fuzzy $\alpha - cuts$ for the algorithmic returns generated by the proposed strategy and compare it with a simple buy-and-hold strategy. We use the hourly Bitcoin price data taken from yahoo! finance from the period of April 03, 2020, 0200 hr to April 1, 2022 2100 hr. The first step is to train an LSTM model on this time series data. The general architecture of an LSTM model is given in Figure 4.3. We use the *PyTorch* framework to develop the prediction model. The hourly closing price is first scaled using a min-max scaling so that the values are normalized between -1 and 1. The data is then split into training and testing sets, with the first 80% allocated to training and the remaining 20% to test set.

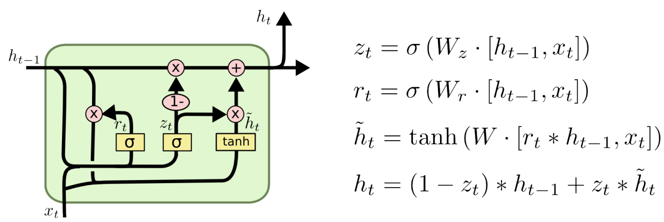

Figure 4.3: General architecture of an LSTM model [2]

After training, the model is used to predict the closing price on the test dataset. The predicted prices are in the range [-1,1] which are then reverse scaled to bring them back to the original scale using the inverse of min-max scaling. Based on the predicted values, a price return prediction is calculated for a timestamp by calculating a simple return ($R_t = Pred(t+1) - Pred(t)$). Based on this predicted return, we calculate the buy/ sell decision for each hour. We set 2 thresholds – buy threshold and sell threshold. The buy threshold is the value we set such that if for any hour the predicted return is greater than this threshold, we buy a bitcoin. While the short selling threshold is reverse of it, i.e., if the predicted loss is greater than this threshold, we short-sell 1 bitcoin. Otherwise, if the predicted return is in between the range $[sell - threshold, buy - threshold]$, then we hold the stock.

Based on this strategy, we calculate the final return by executing this strategy on the test set and calculating the total payoff. We test different thresholds for buy (in range [50, 500) by increments of 10) and sell ((-500, 50] by an increment of 10) to calculate the overall payoff for test data. We see that on average, the total payoff from different combinations of threshold is positive for 5 different runs (the model is retrained in each of these run). Also, some thresholds consistently give high profits. Based on different runs, we can choose the optimal threshold by looking at best average return for different runs.

4.4 Implementation Details

The trend prediction model for the algorithmic trading strategy is implemented in *python* programming language. The data pre-processing is accomplished using the *Pandas* library of python, which provides the construct called *Dataframes* allowing numerous data manipulations. The reshaping of data to bring lagged values of time series in each row input, and splitting the data matrix into training and testing set, is accomplished using the *Numpy* linear algebra library of python. The python library *Scikit-learn* is used for data normalization required for model fitting, and the subsequent denormalization of model predictions. The LSTM model is configured and trained using *PyTorch* library, which is specialized for ease and efficiency of training of Neural Network models.All the experiments were run on a commodity hardware *(11th Gen Intel(R) Core(TM) i5-1135G7 @2.40 Ghz processor with 16 GB of RAM)* with a *64-bit Windows 10* Operating System running over it.

Our LSTM model consists of a single input layer with 20 features, which correspond to the lag period of 20 hours. This is followed by 2 hidden LSTM layers, each with 32 units, followed by a linear output layer. We use the *mean squared error (MSE)* with mean reduction as a loss function for the network training. The *ADAM* optimization algorithm, which is an extension of *stochastic gradient descent,* was used to optimize the network weights. We used 30 epochs for model training in our experiments.

For the risk forecasting models a similar implementation was done in R programing language using its *dplyr* package, which is part of the extentsive *Tidyverse* data science ecosystem.

4.5 Evaluation Metrics

Our proposed trading model for algorithmic trading consists of two separate stages, each with its own evaluation metric. The first stage is the training of LSTM model for prediction of one step ahead price based on lag size of 20 hours. We use the MSE for prediction on the test dataset as evaluation criteria for this stage. In order to calculate the MSE, we first infer the price predictions on test dataset by feeding the test inputs to the trained LSTM model. The predicted price is then compared with the actual price at that hour to calculate the MSE as follows.

$$\text{MSE} = 1/n \sum_{i=1}^{n_{\text{test}}} (y_i - \hat{y}_i)^2 \tag{4.1}$$

The evaluation of risk forecasting models is also done using fuzzy interval widths described above.

The next stage of our trading model is to use the LSTM predictions for making buy/ sell decisions at each hour. This is done by setting two thresholds, the upper one (T_b) to make a buy decision and the lower one (T_s) to make a short-selling decision. For a given pair of the thresholds, $< T_b, T_s >$, we calculate the actual returns (denominated in USD) that would be generated on test data if one unit of Bitcoin was bought when the predicted hourly return exceeds the T_b or sold short if it falls below the T_s.

Our strategy integrated with the fuzzy approach and evaluation metrics could prove advantageous for practitioners in improving profits by doing a finer search for T_b and T_s, which is explained in section 5.4.1.

4.6 Dataset

Figure 4.4: Bitcoin prices for the period of study (2020-04-03 02:00:00 to 2022-04-01 21:00:00)

The algorithmic trading model solely relies on the time series of historical closing prices of Bitcoin captured every hour. We get this data from *Yahoo! finance* API using the python package *yfinance*. The API has a limit on the permissible time interval for fetching the data, as it stores only the last 730 days of hourly closing prices. Due to this limit, the period of our primary study ranges from 2020-04-03 02:00:00 to 2022-04-01 21:00:00. A plot showing the Bitcoin price variation for our study is also given (see figure 4.4). We divide this data into training and testing sets, with the initial 80 percent of the time series used for training and the recent 20 percent used for testing.

As a standard practice for training neural network models, the data has to be normalized to fit into a certain fixed interval. We use the *min-max normalization* technique to scale the closing prices in the range of $[-1, 1]$. The *MinMaxScaler* function available in *Scikit-learn* python library is used for data normalization.

For the risk forecasting studies, we use a daily dataset downloaded using the R package *quantmod* using yahoo finance as the source. For the daily stock and cryptocurrency data, we have the data available for a longer period and we choose the 1000 days observations for our studies in Risk forecasting.

Chapter 5

Experiment Results

We implemented our models and experiments using the methods described in the previous chapter (Chapter 4). In this chapter, we collect the results of these studies and discuss them thoroughly.

5.1 Fuzzy Risk Forecasting of Returns on Stocks and Cryptocurrencies

This section depicts the results of the study carried out to study the effectiveness of data-driven and neuro volatility models in risk forecasting of returns for various stocks and cryptocurrencies.

5.1.1 Summary Statistics

Fig. 5.1 and Fig. 5.2 show the density plots of daily log-returns of stocks and cryptocurrencies. Based on the fitted normal curve, it is clear that stocks/indexes

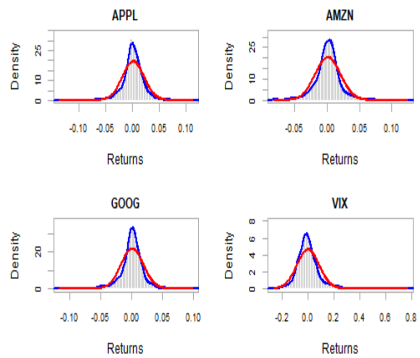

Figure 5.1: Density Plots of Log Returns for Stocks. The blue curve is smoothed density function for the returns, while the red curve is a Gaussian distribution fitted over the empirical distribution

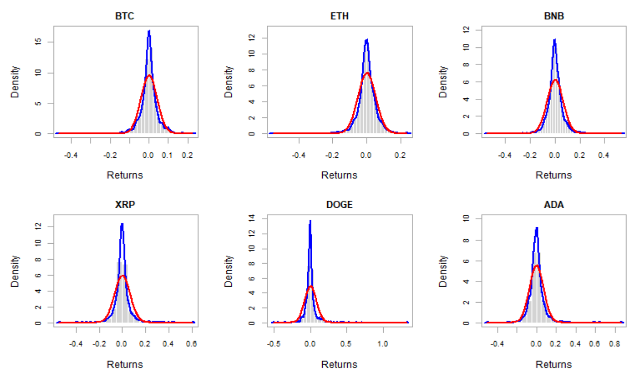

Figure 5.2: Density Plots of Log Returns for Cryptocurrencies. The blue curve is smoothed density function for the returns, while the red curve is a Gaussian distribution fitted over the empirical distribution

and cryptocurrency returns are more peaked at the center with a fatter tail than the normal curve. Although Fig 5.1 and Fig. 5.2 indicate distributions symmetric at a certain level, in general, log-returns of stocks and cryptocurrency have asymmetric distributions.

Table 5.1 summarizes the basic descriptive statistics of the daily log returns. Note that we do not make any assumptions on the distribution such as Gaussianity. Except for VIX and XRP, both the average and the standard deviation of the log returns for cryptocurrencies are high compared to stocks/indexes. It is important to note that even though VIX indicates the lowest average log returns, its standard deviation is the highest among the stocks. VIX has the smallest Sharpe ratio value among the stocks and cryptocurrencies. Thus, VIX gives a better risk-return trade-off compared to other stocks and cryptocurrencies. Among the cryptocurrencies, XRP has the lowest Sharpe ratio, and BNB has the highest. However, the Sharpe ratio of XRP is larger than VIX, and the Sharpe ratio of BNB is smaller than APPL, which has the highest Sharpe ratio among stocks. For Cryptocurrencies and stocks, kurtosis is positive, and it indicates data are heavy-tailed relative to a normal distribution. Nonetheless, compared to cryptocurrencies, kurtosis values for stocks are smaller. Among the stocks, FB has the highest kurtosis, which is 11.34. For APPL, FB, GOOG, BTC, and ETH, returns are negatively skewed, whereas returns of AMZN, BNB, XRP, DOGE, and ADA are positively skewed. It is important to note that the Skewness of the AMZN is the smallest among all the stocks and cryptocurrencies.

Table 5.1: Descriptive statistics of log returns

	Mean	SD	Sharpe	Kurtosis	Skewness
APPL	0.14%	2.03%	0.07	6.33	-0.36
AMZN	0.13%	1.97%	0.06	3.84	0.07
FB	0.07%	2.25%	0.03	11.34	-0.99
GOOG	0.11%	1.81%	0.06	5.52	-0.27
VIX	0.06%	8.56%	0.01	8.87	1.60
BTC	0.17%	4.18%	0.04	12.33	-0.86
ETH	0.17%	5.26%	0.03	10.49	-1.00
BNB	0.39%	6.39%	0.06	12.88	0.37
XRP	0.10%	6.76%	0.01	15.60	0.85
DOGE	0.35%	8.15%	0.04	56.19	3.87
ADA	0.27%	7.25%	0.04	22.14	1.90

5.1.2 VaR and ES Using Traditional Models

We report VaR forecasts using volatility forecasting models including HS, MA, EWMA, Garch, tGarch for all the stocks and cryptocurrencies. We also validate our VaR and ES forecasts against the corresponding forecasts reported in `www.ExtremeRisk.org` (Fig. 5.3, Fig. 5.4).

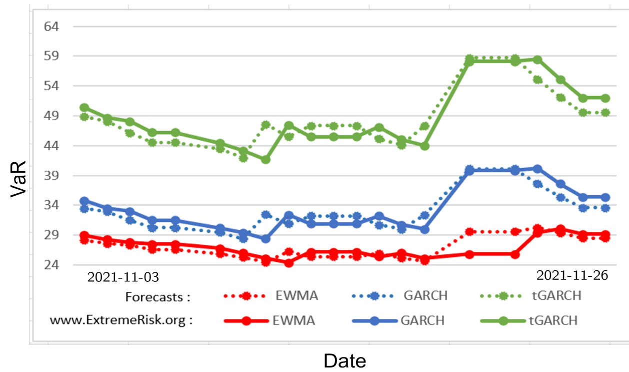

Figure 5.3: VaR Forecasts for Apple (2021-11-03 to 2021-11-26) using different models. The solid line are forecast using traditional model (from ExtremeRisk portal, while data-driven forecasts are in dotted lines.

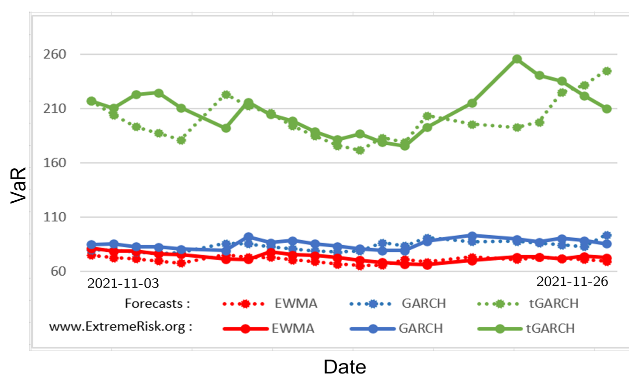

Figure 5.4: VaR Forecasts for Bitcoin (2021-11-03 to 2021-11-26) using different models. The solid line are forecast using traditional model (from ExtremeRisk portal, while data-driven forecasts are in dotted lines.

Forecasts of VaR and ES are reported in Table 5.2 and Table 5.3. Except for APPL and VIX, model risks under VaR for stocks are less compared to cryptocurrencies. Among cryptocurrencies, under VaR, XRP has the highest models risk, and ETH has the lowest. In contrast, DOGE has the highest, and BNB has the lowest model risk under ES. APPL has the highest models risk under VaR and ES among stocks. It is also important to point out that the model risk of APPL is higher than ETH, BNB, and ADA under VaR, and under ES, it is higher than BTC, ETH, BNB, and ADA. Except for APPL, model risks for stocks under both VaR and ES vary between 1.5 and 2.0. However, for cryptocurrencies, model risks are greater than two under all the volatility forecasting models. This is a clear indication of a need of using modified and improved forecast models for both stocks and cryptocurrencies. Table 5.2 and 5.3 also report degrees of freedom for each stock and cryptocurrency. Our results highlight that even though some studies use fixed degrees of freedom, this may not be appropriate for all data.

Table 5.2: Forecasts of VaR and Model Risk

Stock	HS	MA	EWMA	Garch	tGarch	df	MR
APPL	59.88	47.98	28.54	33.60	49.54	4.95	2.10
AMZN	56.00	45.58	40.58	42.03	61.33	5.15	1.51
FB	65.18	53.05	42.00	45.02	69.86	4.08	1.66
GOOG	51.90	42.67	31.83	32.74	52.63	3.84	1.65
VIX	184.18	201.96	120.14	158.23	233.96	4.29	1.95
BTC	108.02	93.25	69.45	93.63	244.92	2.52	3.53
ETH	142.85	118.46	90.19	111.62	200.29	3.49	2.22
BNB	146.76	135.84	103.92	113.53	188.94	3.72	1.82
XRP	169.04	145.64	81.56	176.25	508.62	2.30	6.24
DOGE	196.25	199.21	104.43	150.89	396.66	2.33	3.80
ADA	134.35	136.77	84.16	122.81	210.10	3.70	2.50

Table 5.3: Forecasts of ES and Model Risk

Stock	HS	MA	EWMA	Garch	tGarch	df	MR
APPL	84.38	54.97	32.69	38.49	50.71	4.95	2.58
AMZN	68.66	52.22	46.50	48.15	63.11	5.15	1.48
FB	96.82	60.77	48.11	51.58	69.14	4.08	2.01
GOOG	68.67	48.89	36.46	37.51	51.37	3.84	1.88
VIX	211.39	231.38	137.64	181.28	233.90	4.29	1.70
BTC	165.53	106.83	79.56	107.26	188.48	2.52	2.37
ETH	220.90	135.72	103.32	127.88	190.15	3.49	2.14
BNB	248.91	155.62	119.06	130.06	182.81	3.72	2.09
XRP	279.00	166.85	93.44	201.92	328.29	2.30	3.51
DOGE	307.90	228.23	119.64	172.87	265.01	2.33	2.57
ADA	214.63	156.69	96.42	140.70	202.94	3.70	2.23

5.1.3 Data-Driven Regularized Risk Forecasts

In this section, we report *VaR* and *ES* using data-driven regularized risk forecasting methods. Reported model risk is based on four forecasting models, and the superiority of the forecast models can be observed through model risk. Forecasting models include normal Garch, tGarch, data-driven generalized EWMA, and data-driven EWMA, and summarized results are given in Table 5.4. Observe that compared to regular models reported in the previous section, data-driven models provide more stable VaR and ES forecasts in general. Also, further improvement of the risk forecasting can be noted due to regularization as VaR and ES forecasts get closer to 1.

5.1.4 Comparison of Fuzzy Forecast Intervals

All the experiments were executed on a computer with an Intel Core(TM) i5-8265U CPU, running 4 Cores (8 logical cores) at a clock speed of 1.60GHz and 8GB

Table 5.4: Model Risk for VaR and ES Using DDRRF

	Non-Regularized		Regularized	
	VaR	ES	VaR	ES
APPL	2.09	1.72	1.82	1.53
AMZN	1.07	1.24	1.09	1.37
FB	2.30	1.81	2.17	1.79
GOOG	1.52	1.51	1.17	1.37
VIX	1.45	1.28	1.36	1.20
BTC	1.18	1.60	1.24	1.63
ETH	1.17	1.48	1.23	1.56
BNB	1.80	1.45	1.80	1.45
XRP	1.17	1.72	1.25	1.72
DOGE	1.20	1.77	1.24	1.77
ADA	1.11	1.47	1.19	1.58

of RAM on the Windows 11 operating system. First, we investigate computation time for both Data-Driven Volatility Forecast (DDVF) and neuro volatility Forecasts (NVF) for stocks and cryptocurrencies. It can be seen from Tables 5.5 and 5.6 that NVF takes more computation time compared to DDVF. This is due to maintaining a higher level of accuracy that results from NVF, which uses multiple hidden layers and nodes. Thus, it is important to further investigate if the accuracy of the risk measures could be improved using DDVF and NVF for stocks and cryptocurrencies.

Table 5.5: Computation Time (seconds) for Stocks

	AAPL	AMZN	FB	GOOG	VIX
DDVF	0.21	0.4	0.47	0.17	0.44
NVF	181.03	139.98	232.36	172.95	225.26

Table 5.6: Computation Time (seconds) for Cryptocurrencies

	BTC	ETH	BNB	XRP	DOGE	ADA
DDVF	0.46	0.23	0.52	0.44	0.67	0.48
NVF	419.01	191.19	350.28	272.78	1535.5	283.64

As we develop α-cuts using DDVF and NVF to calculate VaR and ES, there are several cases to be investigated depending on values of m, n, and α (See 3.2.4 for more details). We set m to be 0.25 and n to be 0.25, 0.75, and 1.0, and for α we consider ten equally spaced values within its parameter space $[0, 1]$. Forecasts of VaR and ES interval widths for APPL, BTC, and ADA are reported in Table 5.7, Table 5.8, and Table 5.9 respectively. Note that for APPL, VaR interval widths are smaller using NVF for $n = 0.25$, $n = 0.75$, and $n = 1.0$, whereas for BTC, VaR interval widths using DDVF are smaller. Except for ADA, for all the cryptocurrencies, VaR interval widths using DDVF are smaller, and for stocks, VaR interval widths using NVF are smaller. Thus, we can conclude with one exception that VaR forecasts using DDVF are much more stable for cryptocurrencies, and VaR forecasts using NVF are more stable for stocks. It is important to investigate interval widths when m changes through its parameter space. Our experimental results suggest that except for ADA VaR interval widths using DDVF for cryptocurrencies are smaller compared to VaR interval widths using NVF considering $m = 0.75$ and $m = 1.0$. The results also suggest that VaR interval widths using NVF for stocks are smaller compared to VaR interval widths using DDVF when $m = 0.75$ and $m = 1.0$. Further, the same conclusions can be drawn for ES interval widths.

Thus, the narrowest α-cuts of the annualized volatility are provided using the data-driven volatility forecast for cryptocurrencies, whereas for regular stocks, the narrowest α-cuts of the annualized volatility are provided using the neural volatility forecast. This hold when n and m (tuning parameters) be chosen in between 0 and 1.

Table 5.7: α-cuts of VaR Interval Widths for Apple using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
0	240.08	240.08	240.08	205.41	205.41	205.41
0.1	240.07	237.68	234.93	205.40	200.99	195.90
0.2	239.89	233.95	229.68	205.21	194.24	186.35
0.3	239.13	229.20	224.10	204.38	186.06	176.63
0.4	237.06	223.21	217.79	202.14	176.56	166.55
0.5	232.71	215.49	210.18	197.44	165.65	155.85
0.6	224.79	205.41	200.57	188.87	153.09	144.16
0.7	211.75	192.11	188.07	174.78	138.51	131.05
0.8	191.75	174.60	171.65	153.15	121.49	116.03
0.9	162.66	151.70	150.10	121.70	101.46	98.51
1.0	122.08	122.08	122.08	77.82	77.82	77.82

Table 5.8: α-cuts of VaR Interval Widths for Bitcoin using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
0	655.12	655.12	655.12	1130.77	1130.77	1130.77
0.1	655.08	645.32	634.03	1130.70	1106.57	1078.66
0.2	654.54	630.23	612.72	1129.61	1069.51	1026.26
0.3	652.17	611.55	590.65	1124.88	1024.48	972.82
0.4	645.80	589.09	566.90	1112.15	971.96	917.12
0.5	632.37	561.91	540.17	1085.32	911.14	857.41
0.6	607.95	528.61	508.82	1036.51	840.40	791.47
0.7	567.73	487.35	470.81	956.15	757.45	716.57
0.8	506.04	435.85	423.77	832.87	659.37	629.50
0.9	416.33	371.47	364.93	653.59	542.71	526.53
1.0	291.16	291.16	291.16	403.46	403.46	403.46

5.2 Risk Forecasting for High-Frequency Intraday Data

We also extended the study on high-frequency intraday minute-frequency price data. We use the intraday data for stocks from the Alpha Vantage [56] platform, while for the Cryptocurrencies intraday data, we use the Cryptocompare platform [57]. We

Table 5.9: α-cuts of VaR Interval Widths for Cardano using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
0	1907.15	1907.15	1907.15	1842.45	1842.45	1842.45
0.1	1907.08	1888.80	1867.66	1842.31	1786.29	1721.48
0.2	1906.00	1860.48	1827.72	1840.14	1700.61	1600.17
0.3	1901.32	1825.28	1786.15	1830.77	1597.65	1477.70
0.4	1888.72	1782.54	1741.00	1805.52	1480.01	1352.67
0.5	1862.15	1730.24	1689.54	1752.28	1347.88	1223.11
0.6	1813.85	1665.32	1628.26	1655.48	1200.14	1086.52
0.7	1734.29	1583.81	1552.85	1496.06	1034.71	939.80
0.8	1612.27	1480.87	1458.24	1251.53	848.69	779.33
0.9	1434.81	1350.83	1338.58	895.91	638.47	600.90
1.0	1187.22	1187.22	1187.22	399.77	399.77	399.77

use a window size of 120 minutes for the rolling forecasts of cryptocurrencies due to their 24-hour trading data, while for stocks we use a window size of 90 minutes because data is available only for the extended trading hours (4 a.m. ET to 8 p.m. ET). In order to account for the variability between different days, we run the experiments for October 3, 2022 (Monday), October 5, 2022 (Wednesday), and October 7, 2022 (Friday) which are the opening, middle, and closing days of the week for stock markets.

Unlike the previous study on daily frequency data, no definitive conclusions could be drawn about the relative effectiveness of the neuro volatility model and data-driven volatility model for this extended study. We saw that for both stocks and Cryptocurrencies, data-driven model can give better forecasts on one day, while the neuro volatility model could prove better on other days.

For AAPL, the data-driven forecasting model gives better results on all three days of study as can be seen in Table 5.10, and the plot of the forecasts for October 3 is presented in Figure 5.5. For AMZN, the data-driven forecasting model gives better results on October 7, while the neuro volatility model gives better results on October 3 and October 5. The fuzzy interval widths for AMZN are summarized in Table 5.11, while a forecast plot is presented in Figure 5.6. For GOOGL, the data-driven model gives better forecasts on October 5 and October 7, while the neuro volatility model gives a slightly better forecast on October 3. Some of the results for GOOGL are summarized in Table 5.12, and plotted in Figure 5.7. For META, neuro volatility model gives better results on all three days (see Table 5.13 and Figure5.8 for results). For the CBOE Volatility Index (VIX), we found that data-driven model works better on October 7, and neuro volatility model works better on October 5. For October 3, we see a variation in performance for different α values. The data-driven model gives narrower fuzzy intervals, which correspond to a better forecast, for $\alpha \in [0, 0.7]$, while neuro volatility model gives narrower fuzzy intervals for $\alpha \in (0.7, 1]$. The results are presented in Table 5.14 and Figure 5.9.

Similar results are observed for cryptocurrency assets, as neither of the data-driven or neuro volatility models performs consistently better for all days on all assets. We see that for both Bitcoin and Ethereum, the data-driven model gives better results on October 3, and October 5, while the neuro volatility model has better results for October 7. For Cardano, the data-driven model gives better forecasts on October 3, while for other days neuro volatility model gives better forecasts. Some of the results for Cryptocurrencies are given in Table 5.15 (and Figure 5.10) for Bitcoin, Table 5.16 (and Figure 5.11) for Ethereum, and Table 5.17 (and Figure 5.12) for Cardano.

Further experiments with large volumes of high-frequency intraday data are required to reach any valuable conclusion. However, we leave this as future work.

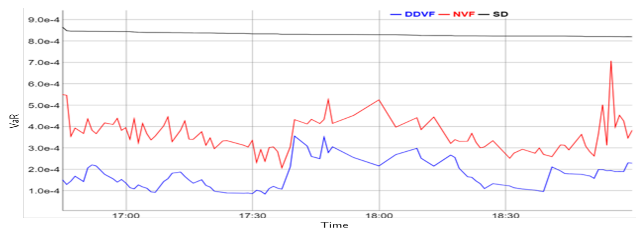

Figure 5.5: Intraday VaR forecasts of AAPL using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 3, 2022

Table 5.10: α-cuts of VaR Interval Widths for AAPL intra-day minute frequency data (dated October 3, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
0	12.22	12.22	12.22	15.10	15.10	15.10
0.1	12.22	11.90	11.53	15.10	14.91	14.70
0.2	12.20	11.41	10.84	15.09	14.62	14.28
0.3	12.15	10.83	10.15	15.04	14.26	13.85
0.4	12.01	10.16	9.44	14.91	13.82	13.39
0.5	11.71	9.41	8.70	14.63	13.27	12.85
0.6	11.16	8.57	7.93	14.13	12.60	12.22
0.7	10.26	7.64	7.10	13.30	11.75	11.43
0.8	8.87	6.59	6.19	12.03	10.67	10.44
0.9	6.86	5.40	5.19	10.18	9.31	9.19
1.0	4.05	4.05	4.05	7.60	7.60	7.60

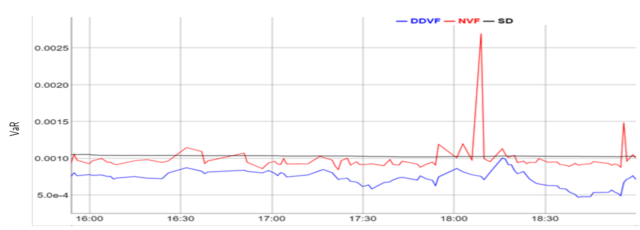

Figure 5.6: Intraday VaR forecasts of AMZN using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 5, 2022

5.3 Risk Forecasting of Algo Returns

The results reported in this section correspond to the experimental methods suggested in Chapter 4.2 for the risk forecasting of algorithmic returns A_t generated by

Table 5.11: α-cuts of VaR Interval Widths for AMZN intra-day minute frequency data (dated October 5, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	16.42	16.42	16.42	11.20	11.20	11.20
2	16.42	16.28	16.1.0	11.20	10.89	10.53
3	16.41	16.04	15.78	11.19	10.41	9.85
4	16.34	15.73	15.42	11.14	9.84	9.17
5	16.15	15.30	14.97	11.00	9.18	8.47
6	15.76	14.71	14.38	10.70	8.44	7.75
7	15.05	13.86	13.57	10.16	7.62	6.98
8	13.88	12.68	12.43	9.26	6.69	6.16
9	12.09	11.04	10.86	7.89	5.64	5.26
10	9.48	8.81	8.71	5.90	4.46	4.25
1.0	5.84	5.84	5.84	3.1.0	3.1.0	3.11

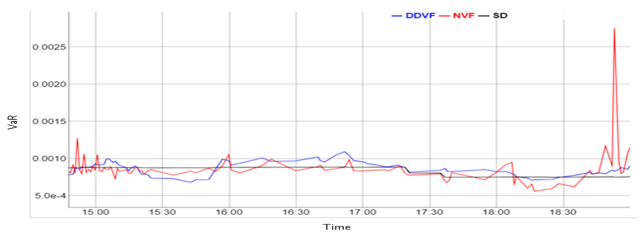

Figure 5.7: Intraday VaR forecasts of GOOGL using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 5, 2022

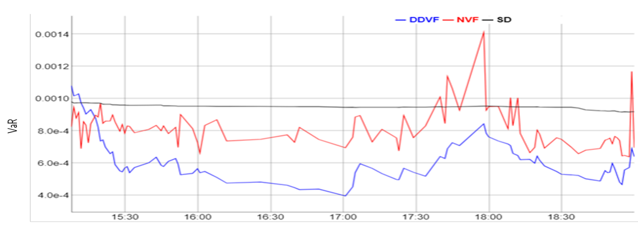

Figure 5.8: Intraday VaR forecasts of META using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 5, 2022

the SMA crossover strategy. The summary statistics of A_t for all the assets are listed in Table 5.18 . For stocks/indexes, the sample sign correlation of algorithmic returns ($\hat{\rho}_A$) vary between 0.64 and 0.70. For all the cryptocurrencies, the $\hat{\rho}_A$ is less than 0.71

Table 5.12: α-cuts of VaR Interval Widths for GOOGL intra-day minute frequency data (dated October 3, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	24.84	24.84	24.84	23.53	23.53	23.53
2	24.84	24.48	24.07	23.53	23.02	22.42
3	24.81	23.93	23.30	23.51	22.22	21.30
4	24.72	23.25	22.49	23.41	21.26	20.15
5	24.46	22.41	21.61	23.14	20.13	18.96
6	23.92	21.37	20.59	22.56	18.83	17.68
7	22.94	20.07	19.36	21.52	17.32	16.28
8	21.33	18.42	17.82	19.81	15.55	14.68
9	18.85	16.31	15.87	17.18	13.46	12.82
10	15.25	13.62	13.39	13.35	10.98	10.63
1.0	10.22	10.22	10.22	8.02	8.02	8.02

Table 5.13: α-cuts of VaR Interval Widths for META intra-day minute frequency data (dated October 5, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	22.29	22.29	22.29	15.37	15.37	15.37
2	22.29	21.73	21.09	15.37	15.08	14.74
3	22.26	20.88	19.89	15.36	14.63	14.11
4	22.16	19.86	18.67	15.30	14.08	13.46
5	21.89	18.67	17.41	15.13	13.43	12.77
6	21.32	17.32	16.09	14.77	12.66	12.01
7	20.28	15.78	14.65	14.1.0	11.74	11.15
8	18.57	14.00	13.07	13.04	10.63	10.14
9	15.94	11.96	11.27	11.39	9.29	8.93
10	12.12	9.57	9.20	8.99	7.64	7.45
1.0	6.79	6.79	6.79	5.64	5.64	5.64

and greater than 0.36. Except for ADA, for all the assets, the algorithmic returns have a t-distribution with degrees of freedom (d.f) less than 4. However, note that the corresponding d.f. for ADA is very close to 4. Moreover, the absolute algorithmic returns $|A_t|$ are significantly autocorrelated for all the assets, which indicate volatility clustering. Figure 5.13 and Figure 5.14 visualize volatility clustering plots for AAPl

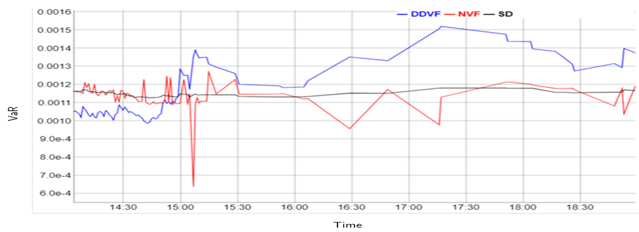

Figure 5.9: Intraday VaR forecasts of VIX using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 3, 2022

Table 5.14: α-cuts of VaR Interval Widths for META intra-day minute frequency data (dated October 3, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	108.55	108.55	108.55	176.66	176.66	176.66
2	108.54	104.71	100.28	176.64	169.50	161.24
3	108.41	98.87	92.01	176.40	158.61	145.81
4	107.83	91.89	83.69	175.35	145.64	130.34
5	106.25	84.00	75.30	172.54	131.04	114.81
6	102.93	75.29	66.76	166.61	115.05	99.15
7	96.90	65.78	58.01	155.82	97.77	83.29
8	86.97	55.43	48.95	138.06	79.24	67.14
9	71.73	44.20	39.45	110.81	59.45	50.61
10	49.58	31.98	29.41	71.18	38.36	33.58
1.0	18.66	18.66	18.66	15.90	15.90	15.90

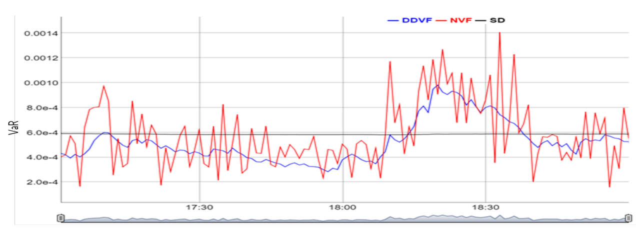

Figure 5.10: Intraday VaR forecasts of Bitcoin using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 3, 2022

and BTC, respectively. There is a clear indication that within the study periods for AAPL and BTC, there are periods with low volatility and periods with high volatility. Volatility is high during March and April in 2021 (peak period of COVID-19

Table 5.15: α-cuts of VaR Interval Widths for Bitcoin intra-day minute frequency data (dated October 3, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	25.35	25.35	25.35	37.70	37.70	37.70
2	25.35	24.68	23.91	37.70	37.01	36.20
3	25.32	23.66	22.46	37.67	35.94	34.69
4	25.20	22.43	21.00	37.52	34.63	33.15
5	24.88	21.01	19.49	37.13	33.10	31.52
6	24.21	19.40	17.91	36.30	31.29	29.75
7	23.00	17.58	16.22	34.80	29.16	27.75
8	20.99	15.50	14.37	32.33	26.61	25.44
9	17.91	13.12	12.29	28.54	23.54	22.69
10	13.44	10.37	9.93	23.02	19.83	19.36
11	7.19	7.19	7.19	15.32	15.32	15.32

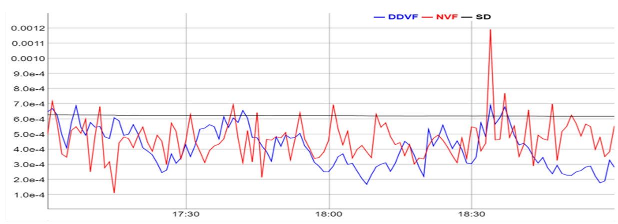

Figure 5.11: Intraday VaR forecasts of Ethereum using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 7, 2022

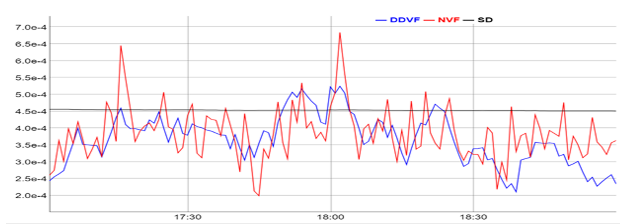

Figure 5.12: Intraday VaR forecasts of Cardano using data-driven model, neuro volatility model and Historical sample Standard Deviation on October 7, 2022

pandemic) for both assets.

Table 5.19 summarizes α-cuts of Annualized Sharpe Ratio (ASR) estimates using risk as DDVE. DDVE are calculated using equation 3.14 following computation of

Table 5.16: α-cuts of VaR Interval Widths for Ethereum intra-day minute frequency data (dated October 7, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	17.85	17.85	17.85	16.89	16.89	16.89
2	17.85	17.61	17.34	16.89	16.57	16.20
3	17.84	17.24	16.81	16.88	16.08	15.51
4	17.79	16.79	16.27	16.81	15.48	14.80
5	17.63	16.24	15.69	16.63	14.77	14.04
6	17.32	15.58	15.05	16.24	13.93	13.22
7	16.74	14.79	14.30	15.53	12.93	12.29
8	15.79	13.81	13.41	14.37	11.74	11.20
9	14.33	12.61	12.31	12.59	10.29	9.89
10	12.22	11.11	10.95	9.99	8.52	8.31
11	9.26	9.26	9.26	6.38	6.38	6.38

Table 5.17: α-cuts of VaR Interval Widths for Cardano intra-day minute frequency data (dated October 7, 2022) using DDVF and NVF ($m = 0.25$)

	DDVF			NVF		
$\alpha \setminus n$	0.25	0.75	1.0	0.25	0.75	1.0
1	11.41	11.41	11.41	10.32	10.32	10.32
2	11.41	11.23	11.02	10.32	10.15	9.96
3	11.39	10.96	10.64	10.31	9.89	9.60
4	11.35	10.61	10.24	10.27	9.58	9.22
5	11.23	10.20	9.80	10.16	9.19	8.81
6	10.96	9.69	9.30	9.93	8.73	8.36
7	10.49	9.06	8.70	9.52	8.16	7.83
8	9.71	8.26	7.96	8.83	7.46	7.18
9	8.51	7.24	7.02	7.79	6.59	6.38
10	6.77	5.96	5.84	6.26	5.50	5.38
11	4.34	4.34	4.34	4.13	4.13	4.13

sample sign correlation of A_t. Table 5.19 captures the ASR and daily algorithmic volatility estimates that are based on the asymptotic variance of the data-driven volatility estimator. Considering all the assets, AAPL has the highest ASR value at 1.40. Among the cryptocurrencies, ADA has the highest ASR, and it is close to the ASR of AAPL. Further, VIX has the lowest ASR among the stocks/indexes, and

Table 5.18: Summary Statistics of Daily algorithmic Returns for All Assets

Assets	Mean	SD	kurtosis	A_t	$\lvert A_t \rvert$	$\lvert A_t^2 \rvert$	$\hat{\rho}_A$	Est df
AAPL	0.0018	0.0193	6.4445	-0.1370	0.2420	0.3245	0.6804	3.4837
AMZN	0.0007	0.0187	4.6384	-0.0308	0.1988	0.1585	0.6965	3.7634
FB	0.0000	0.0210	9.5641	-0.1065	0.1680	0.1266	0.6782	3.4507
GOOG	0.0004	0.0171	5.9019	-0.1290	0.2378	0.2829	0.6769	3.4323
VIX	-0.0023	0.0936	24.2155	-0.0817	0.1832	0.0741	0.6401	3.0288
BTC	0.0022	0.0390	7.7319	-0.0807	0.1306	0.0844	0.6771	3.4353
ETH	0.0028	0.0504	5.7681	-0.0685	0.1127	0.1140	0.6991	3.8163
BNB	0.0028	0.0583	19.1238	-0.0424	0.2495	0.1734	0.6493	3.1121
XRP	0.0017	0.0619	14.0692	-0.0064	0.2685	0.1442	0.6112	2.8208
DOGE	0.0048	0.1207	526.1476	0.0422	0.2245	0.0136	0.3604	2.1590
ADA	0.0044	0.0599	3.8096	-0.0390	0.1586	0.1273	0.7095	4.0604

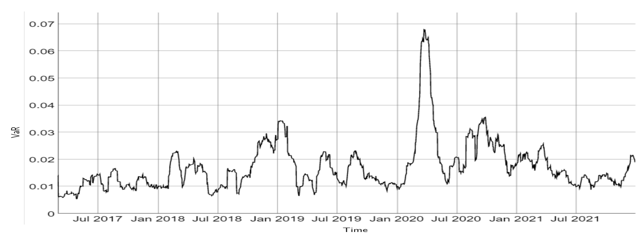

Figure 5.13: Volatility Clustering demonstrated graphically for AAPL. Note that the periods of high volatility are followed the the periods of low volatility.

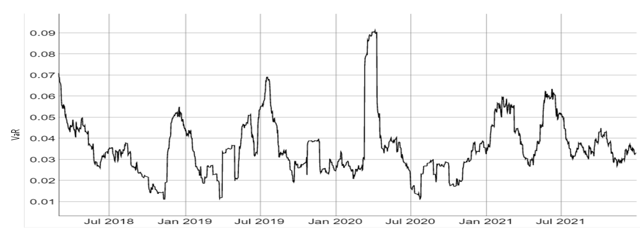

Figure 5.14: Volatility Clustering demonstrated graphically for BTC. Note that the periods of high volatility are followed the the periods of low volatility.

XRP has the lowest among cryptocurrencies.

Similarly, sample standard deviation (s_n), mean absolute deviation (MAD) ($\hat{\rho}_A s_n$), and Value at Risk at $\alpha = 0.05$ ($VaR_{0.05}$) based on t distribution of A_t are used to estimate the daily volatility and annualized SR. Results are summarized in Tables 5.20, 5.21, and 5.22 for all assets. It can be seen that fuzzy α-cut estimates of

Table 5.19: α-CUTs for Annualized SR Using DDVE

Assets	DDVE	0.05-cut of DDVE	ASR	0.05-cut of ASR
AAPL	0.0193	(0.0181, 0.0204)	1.3957	(1.3167, 1.4847)
AMZN	0.0187	(0.0176, 0.0197)	0.5293	(0.5006, 0.5615)
FB	0.0210	(0.0197, 0.0222)	-0.0443	(-0.0418, -0.0472)
GOOG	0.0171	(0.0161, 0.0182)	0.2775	(0.2616, 0.2954)
VIX	0.0935	(0.0873, 0.0998)	-0.4010	(-0.3759, -0.4297)
BTC	0.0389	(0.0367, 0.0411)	1.0417	(0.9862, 1.1038)
ETH	0.0503	(0.0477, 0.0530)	1.0419	(0.9895, 1.1002)
BNB	0.0583	(0.0547, 0.0618)	0.8943	(0.8431, 0.9520)
XRP	0.0618	(0.0577, 0.0660)	0.5200	(0.4873, 0.5574)
DOGE	0.1193	(0.1033, 0.1353)	0.7629	(0.6728, 0.8810)
ADA	0.0598	(0.0567, 0.0629)	1.3778	(1.3104, 1.4525)

annualized SR using DDVE are narrower than that using sample Standard Deviation (SD) and MAD. This means that the volatility forecasts generated by DDVE are more stable and uniform as compared to SD and MAD. MAD has the largest fuzzy α-cut estimates of annualized SR considering all the assets.

Table 5.20: α-CUTs for Annualized SR Using SD

Assets	SD	0.05-cut CI of SD	ASR	0.05-cut of ASR
AAPL	0.0193	(0.0177, 0.0209)	1.3943	(1.2900, 1.5171)
AMZN	0.0187	(0.0173, 0.0200)	0.5290	(0.4936, 0.5699)
FB	0.0210	(0.0190, 0.0230)	-0.0442	(-0.0404, -0.0489)
GOOG	0.0171	(0.0158, 0.0185)	0.2774	(0.2572, 0.3009)
VIX	0.0936	(0.0803, 0.1069)	-0.4007	(-0.3507, -0.4674)
BTC	0.0390	(0.0358, 0.0421)	1.0405	(0.9627, 1.1319)
ETH	0.0504	(0.046, 0.0540)	1.0407	(0.9707, 1.1217)
BNB	0.0583	(0.0514, 0.0653)	0.8930	(0.7980, 1.0136)
XRP	0.0619	(0.0555, 0.0683)	0.5198	(0.4709, 0.5800)
DOGE	0.1207	(0.0489, 0.1926)	0.7540	(0.4727, 1.8623)
ADA	0.0599	(0.0562, 0.0637)	1.3751	(1.2943, 1.4666)

We obtain the rolling forecasts of the daily and annualized SR as per Chapter 4.2. Results are summarized in Table 5.23 and the average RMSEs of the assets are

Table 5.21: α-CUTs for Annualized SR Using MAD

Assets	MAD	0.05-cut of MAD	ASR	0.05-cut of ASR
AAPL	0.0131	(0.0121, 0.0142)	2.0493	(1.8959, 2.2297)
AMZN	0.0130	(0.0121, 0.0139)	0.7595	(0.7087, 0.8182)
FB	0.0143	(0.0129, 0.0156)	-0.0652	(-0.0596, -0.0720)
GOOG	0.0116	(0.0107, 0.0125)	0.4098	(0.3800, 0.4446)
VIX	0.0599	(0.0514, 0.0685)	-0.6261	(-0.5480, -0.7301)
BTC	0.0264	(0.0243, 0.0285)	1.5367	(1.4218, 1.6717)
ETH	0.0352	(0.0327, 0.0378)	1.4887	(1.3885, 1.6045)
BNB	0.0379	(0.0334, 0.0424)	1.3752	(1.2289, 1.5610)
XRP	0.0378	(0.0339, 0.0417)	0.8503	(0.7704, 0.9488)
DOGE	0.0435	(0.0176, 0.0694)	2.0918	(1.3113, 5.1669)
ADA	0.0425	(0.0399, 0.0452)	1.9381	(1.8243, 2.0672)

Table 5.22: α-CUTs for Annualized SR Using $VaR_{0.05}$

Assets	$VaR_{0.05}$	0.05-cut of $VaR_{0.05}$	ASR	0.05-cut of ASR
AAPL	0.0262	(0.0246, 0.0279)	1.0257	(0.9640, 1.0960)
AMZN	0.0270	(0.0254, 0.0286)	0.3654	(0.3451, 0.3883)
FB	0.0304	(0.0285, 0.0322)	-0.0306	(-0.0289, -0.0326)
GOOG	0.0244	(0.0229, 0.0259)	0.1950	(0.1837, 0.2078)
VIX	0.1301	(0.1215, 0.1386)	-0.2883	(-0.2706, -0.3086)
BTC	0.0541	(0.0509, 0.0573)	0.7495	(0.7080, 0.7962)
ETH	0.0722	(0.0683, 0.0762)	0.7258	(0.6879, 0.7681)
BNB	0.0780	(0.0731, 0.0829)	0.6681	(0.6285, 0.7129)
XRP	0.0789	(0.0734, 0.0843)	0.4079	(0.3817, 0.4379)
DOGE	0.0851	(0.0729, 0.0973)	1.0693	(0.9352, 1.2481)
ADA	0.0861	(0.0814, 0.0907)	0.9574	(0.9082, 1.0122)

reported in the second column of Table 5.23. Among all the assets, DOGE has the highest RMSE. BTC has the minimum RMSE among cryptocurrencies. However, note that the RMSE of all the stocks/indexes is smaller than BTC except for VIX. This implies that the algorithmic trading returns (using SMA crossover strategy) generated for stocks have a more reliable ASR forecasts as compared to cryptocurrencies. However, the algorithmic trading returns generated for CBOE VIX, which is a highly volatile index fund, have less reliable ASR estimates than BTC, while DODGE has the least reliable ASR forecast among all assets.

The majority of the research literature only report point forecasts of SR depicted by the red line in Figure 5.15 and 5.16. But, our work also presents the fuzzy α-cuts of the forecast, which captures more realistic uncertainty estimates. Table 5.23 summarizes the results for all the assets, and Figures 5.15 and 5.16 visualize the α-cuts for AAPL and BTC. The fuzzy forecast intervals are highly relevant for this forecasting problem due to the fact that volatility, and by extension Sharpe Ratio, is highly uncertain and exhibits temporal variance.

Table 5.23: SR Fuzzy Forecasts - DDVF

Assets	RMSE	DSR	ASR	0.05-cut of ASR	0.01-cut of ASR
AAPL	0.015	0.097	1.547	(-4.269, 7.363)	(-6.097, 9.191)
AMZN	0.015	0.024	0.376	(-3.957, 4.710)	(-5.319, 6.072)
FB	0.019	-0.003	-0.050	(-4.166, 4.065)	(-5.459, 5.359)
GOOG	0.015	0.029	0.467	(-4.067, 5.002)	(-5.492, 6.427)
VIX	0.074	-0.049	-0.773	(-5.070, 3.524)	(-6.420, 4.874)
BTC	0.038	0.060	1.146	(-3.031, 5.322)	(-4.343, 6.634)
ETH	0.046	0.051	0.981	(-2.940, 4.902)	(-4.172, 6.135)
BNB	0.051	0.019	0.365	(-5.186, 5.917)	(-6.931, 7.661)
XRP	0.062	0.044	0.832	(-2.999, 4.663)	(-4.203, 5.867)
DOGE	0.125	0.054	1.030	(-3.807, 5.868)	(-5.327, 7.388)
ADA	0.053	0.078	1.481	(-2.889, 5.851)	(-4.262, 7.224)

Next, we compute daily and annualized rolling SR forecasts applying the neuro volatility model using the similar window sizes as that of the data-driven rolling forecast. Table 5.24 summarizes the results from the neuro volatility model. Also, daily rolling SR forecasts and their averages for AAPL and BTC are plotted in Figures 5.17 and Figures 5.18 respectively. Fuzzy α-cuts for assets are provided in Table 5.24, and plotted in Figure 5.17 and 5.18. Observations suggest that data-driven volatility forecast (DDVF) provides narrower α-cuts compared to neuro volatility forecasts

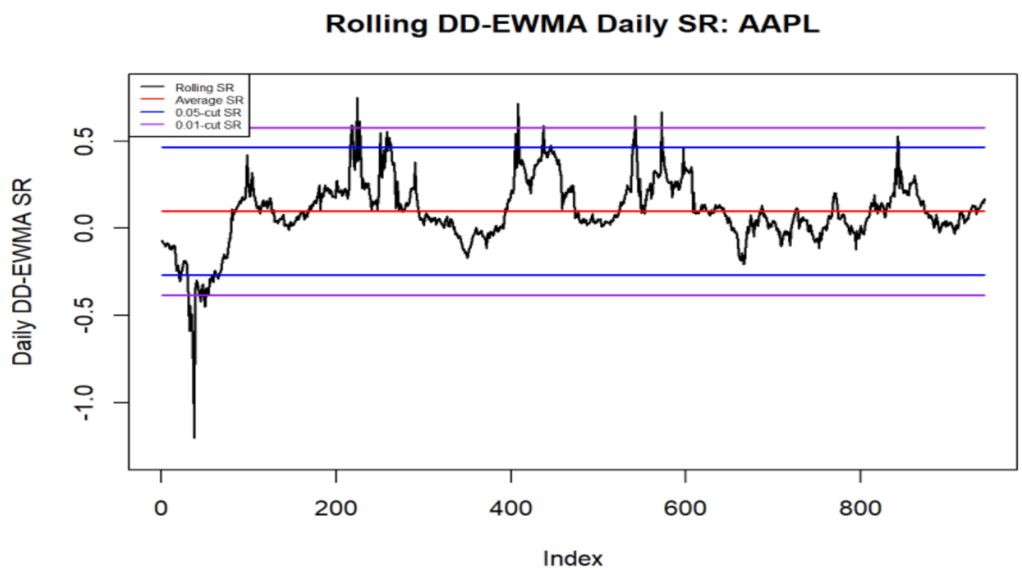

Figure 5.15: Daily rolling SR forecasts using DD-EWMA for AAPL. The plot also provide the average SR in red, and 0.05-α-cut in blue, and 0.01 alpha cut in violet.

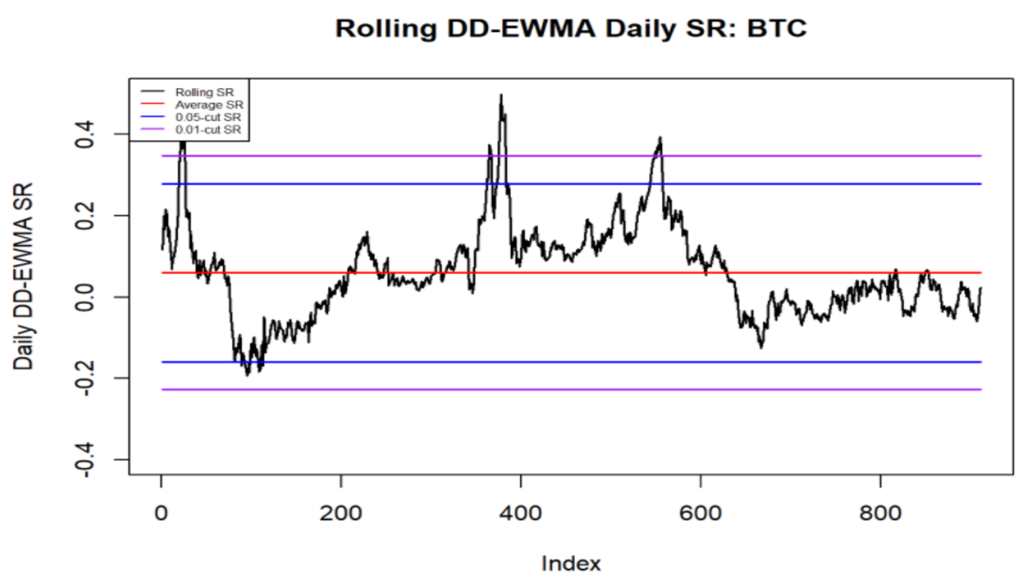

Figure 5.16: Daily rolling SR forecasts using DD-EWMA for BTC. The plot also provide the average SR in red, and 0.05-α-cut in blue, and 0.01 alpha cut in violet.

(NVF). This means that DDVF forecasts are more reliable than NVF forecasts due to their stability.

Computation times for both DDVF and NVF were investigated under both stocks/indexes and cryptocurrencies. NVF takes more computation time compared to DDVF. Thus, it is important to investigate further if the accuracy of the forecasts could be improved using DDVF and NVF for stocks and cryptocurrencies. All the experiments were executed on a computer with an Intel Core(TM) i5-8265U CPU, running 4 Cores (8 logical cores) at a clock speed of 1.60GHz and 8GB of RAM on the Windows 11 operating system.

Table 5.24: SR Fuzzy Forecasts - NVF

Assets	DSR	ASR	0.05-cut of ASR	0.01-cut of ASR
AAPL	0.100	1.581	(-4.420, 7.581)	(-6.305, 9.467)
AMZN	0.025	0.403	(-3.915, 4.721)	(-5.272, 6.078)
FB	-0.006	-0.102	(-3.804, 3.600)	(-4.967, 4.763)
GOOG	0.020	0.319	(-6.300, 6.939)	(-8.380, 9.019)
VIX	-0.044	-0.697	(-4.931, 3.537)	(-6.262, 4.867)
BTC	0.067	1.271	(-4.988, 7.530)	(-6.955, 9.497)
ETH	0.068	1.297	(-10.081, 12.674)	(-13.656, 16.249)
BNB	-0.001	-0.018	(-23.444, 23.407)	(-30.805, 30.768)
XRP	0.051	0.983	(-4.656, 6.623)	(-6.428, 8.395)
DOGE	0.067	1.273	(-16.437, 18.983)	(-22.002, 24.548)
ADA	0.076	1.444	(-5.525, 8.414)	(-7.715, 10.604)

As described in our approach (see Chapter 3.3.2), we use the non-linear adaptive trapezoidal fuzzy numbers to compute forecast intervals. The membership function for such a fuzzy number is parameterized by two additional variables, m and n. We compute the fuzzy intervals using different values of m and n. We consider the following six cases for both the neuro forecast and data-driven forecast : **case1**: $m = 0.25$, $n = 0.25, 0.50, 0.75$, **case2**: $m = 0.50$, $n = 0.25, 0.50, 0.75$, **case3**: $m = 0.75$, $n = 0.25, 0.50, 0.75$, **case4:** $m = 0.50$, $n = 1.0, 2.0, 3.0$, **case5**: $m = 1.0$, $n = 1.0, 2.0, 3.0$, **case6**: $m = 2.0$, $n = 1.0, 2.0, 3.0$). Table 5.25 reports the width of

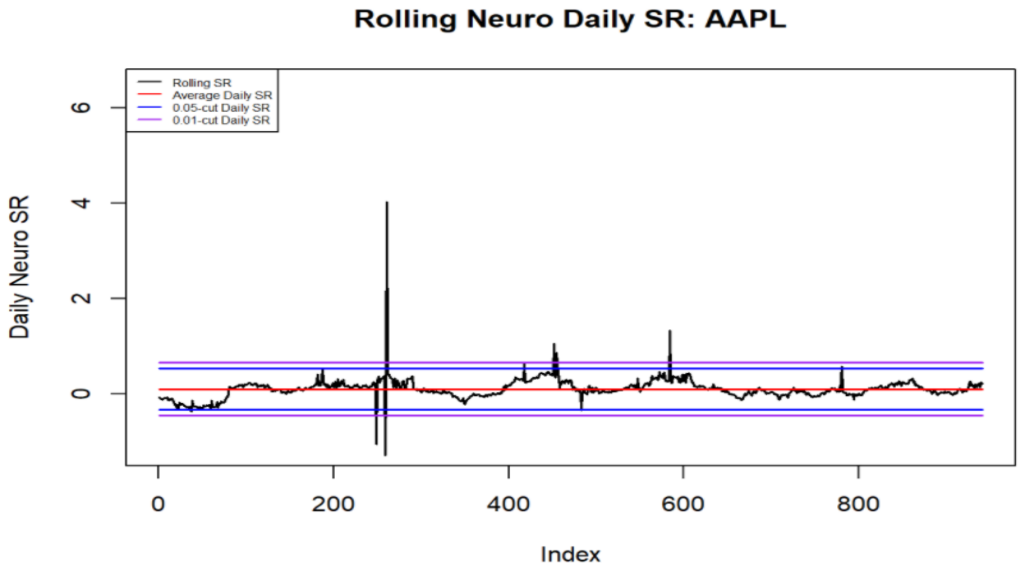

Figure 5.17: Daily rolling SR forecasts using Neuro volatility model for AAPL. The plot also provide the average SR in red, and 0.05-α-cut in blue, and 0.01 alpha cut in violet.

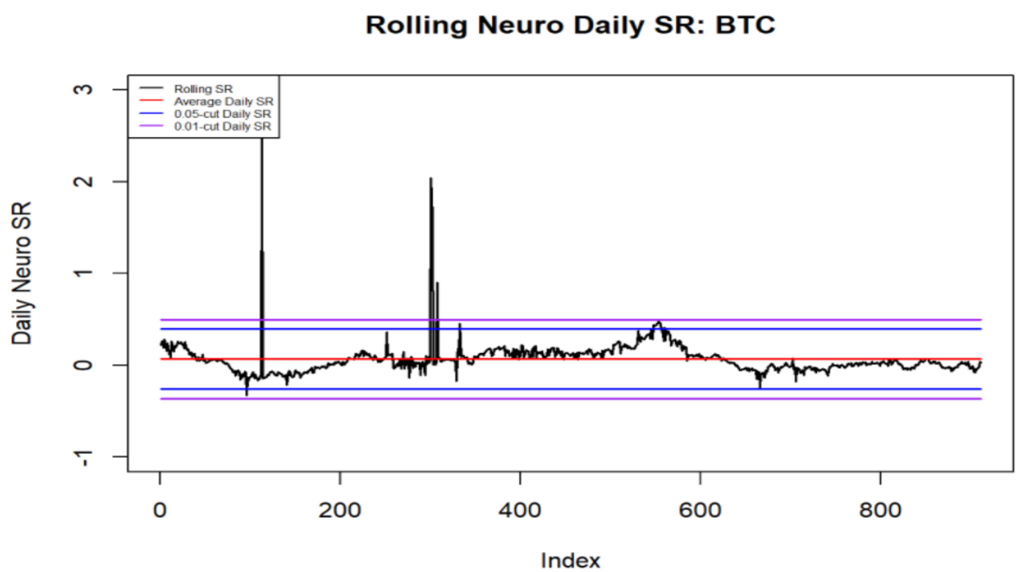

Figure 5.18: Daily rolling SR forecasts using Neuro volatility model for BTC. The plot also provide the average SR in red, and 0.05-α-cut in blue, and 0.01 alpha cut in violet.

fuzzy intervals for BTC using DDVF and NVF when ($m = 0.25$, $n = 0.75$), ($m = 1.0$, $n = 1.0$), and ($m = 1.0$, $n = 2.0$). We found that for all the top six cryptocurrencies, as well as for the stocks AMZN, FB, and GOOG, the fuzzy interval widths computed for the DDVF are narrower than the corresponding intervals obtained for the NVF (Table 5.26). However, for the stock AAPL and the CBOE Volatility index (VIX), the fuzzy intervals obtained for NVF are smaller compared to DDVF (Table 5.27).

Table 5.25: α-cuts of Volatility Interval Widths for BTC using DDVF and NVF

	DDVF			NVF		
	$m = 0.25$	$m = 1.0$	$m = 1.0$	$m = 0.25$	$m = 1.0$	$m = 1.0$
α	$n = 0.75$	$n = 1.0$	$n = 2.0$	$n = 0.75$	$n = 1.0$	$n = 2.0$
0.0	0.0135	0.0135	0.0135	0.6259	0.6259	0.6259
0.1	0.0132	0.0127	0.0113	0.6183	0.5779	0.5426
0.2	0.0128	0.0120	0.0104	0.6063	0.5299	0.4895
0.3	0.0122	0.0112	0.0096	0.5905	0.4819	0.4414
0.4	0.0116	0.0104	0.0089	0.5696	0.4339	0.3959
0.5	0.0109	0.0096	0.0083	0.5412	0.3859	0.3521
0.6	0.0101	0.0089	0.0077	0.5022	0.3379	0.3094
0.7	0.0092	0.0081	0.0072	0.4483	0.2900	0.2676
0.8	0.0082	0.0073	0.0067	0.3749	0.2420	0.2265
0.9	0.0071	0.0065	0.0062	0.2762	0.1940	0.1860
1.0	0.0058	0.0058	0.0058	0.1460	0.1460	0.1460

Thus, the narrowest α-cuts of the annualized volatility are provided using the data-driven volatility forecast for cryptocurrencies, whereas for regular stocks, data-driven or neuro volatility forecasts provide narrowest α-cuts depending on the stocks/indexes. This holds for any two points of n and m (tuning parameters) in their parameter space.

5.4 Proposed Novel Algorithmic Trading Strategy

This section describes the experimental results pertaining to the proposed algorithmic trading strategy for cryptocurrencies, which is based on the price prediction

Table 5.26: α-cuts of Volatility Interval Widths for AMZN using DDVF and NVF

	DDVF			NVF		
	$m = 0.25$	$m = 1.0$	$m = 1.0$	$m = 0.25$	$m = 1.0$	$m = 1.0$
α	$n = 0.75$	$n = 1.0$	$n = 2.0$	$n = 0.75$	$n = 1.0$	$n = 2.0$
0.0	0.1170	0.1170	0.1170	0.1982	0.1982	0.1982
0.1	0.1158	0.1109	0.1054	0.1946	0.1834	0.1667
0.2	0.1140	0.1047	0.0985	0.1890	0.1686	0.1496
0.3	0.1116	0.0986	0.0924	0.1821	0.1539	0.1348
0.4	0.1086	0.0925	0.0867	0.1736	0.1391	0.1212
0.5	0.1047	0.0864	0.0812	0.1631	0.1243	0.1084
0.6	0.0996	0.0803	0.0759	0.1500	0.1096	0.0961
0.7	0.0927	0.0742	0.0708	0.1333	0.0948	0.0843
0.8	0.0836	0.0681	0.0657	0.1120	0.0801	0.0728
0.9	0.0715	0.0620	0.0608	0.0849	0.0653	0.0615
1.0	0.0559	0.0559	0.0559	0.0505	0.0505	0.0505

Table 5.27: α-cuts of Volatility Interval Widths for AAPL using DDVF and NVF

	DDVF			NVF		
	$m = 0.25$	$m = 1.0$	$m = 1.0$	$m = 0.25$	$m = 1.0$	$m = 1.0$
α	$n = 0.75$	$n = 1.0$	$n = 2.0$	$n = 0.75$	$n = 1.0$	$n = 2.0$
0.0	0.2512	0.2512	0.2512	0.1397	0.1397	0.1397
0.1	0.2478	0.2405	0.2250	0.1350	0.1272	0.1053
0.2	0.2427	0.2298	0.2121	0.1278	0.1148	0.0897
0.3	0.2365	0.2192	0.2014	0.1191	0.1024	0.0772
0.4	0.2291	0.2085	0.1918	0.1092	0.0899	0.0663
0.5	0.2205	0.1979	0.1830	0.0980	0.0775	0.0565
0.6	0.2103	0.1872	0.1747	0.0854	0.0650	0.0473
0.7	0.1982	0.1765	0.1667	0.0711	0.0526	0.0387
0.8	0.1836	0.1659	0.1591	0.0549	0.0401	0.0305
0.9	0.1659	0.1552	0.1517	0.0364	0.0277	0.0227
1.0	0.1446	0.1446	0.1446	0.0152	0.0152	0.0152

of LSTM.

5.4.1 LSTM Price Prediction Model

To account for the inherent non-determinism in neural network based models, we trained our LSTM price prediction model five times and took the average MSE. We note the time taken for the LSTM network's training, the average of five runs came

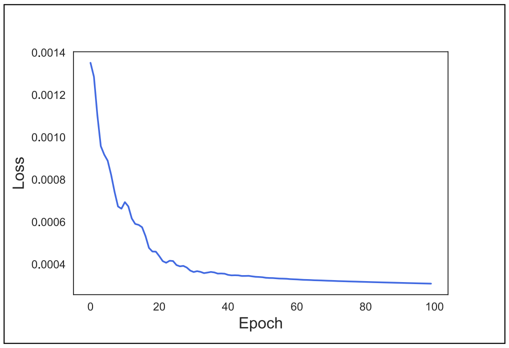

Figure 5.19: Learning curve when LSTM network is trained for 100 epochs. Note that learning curve flats out after around 30 epochs

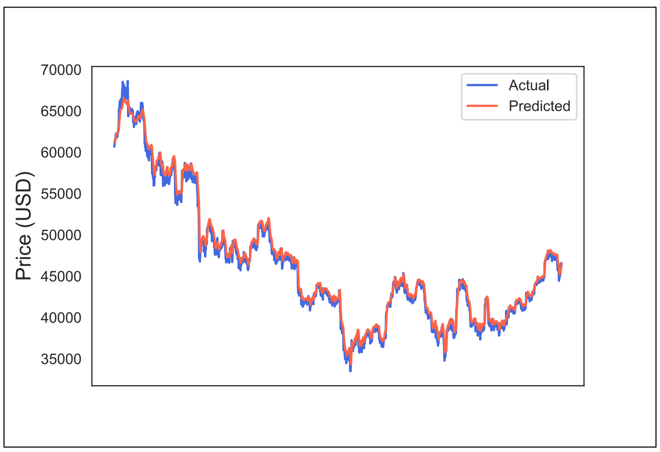

Figure 5.20: Actual vs Forecast prices for test dataset on Bitcoin. Predicted values closely follow the Actuals.

out to be 116.52 seconds. The training of LSTM model was initially done for 100 epochs, but as it can be seen in Figure 5.19, the learning curve becomes flat after around 30 epochs. Hence, to save on computation time, we trained the model for 30 runs in all the subsequent experiments. The trained model is used to forecast one-step ahead price of cryptocurrency, and the forecasts closely follow the actual price movements (see Figure 5.20). The efficacy of the forecasting model is also quantified by MSE for test set, which was recorded at 0.00147 for the average of five runs. We then built our trading strategy, described in Chapter 4.3 over this trained LSTM forecasting model.

5.4.2 Trading Strategy Hyperparameter Tuning and Results for Bitcoin

The trading strategy we proposed is characterized by two key hyperparameters: threshold for buy T_b and the threshold for short-sell T_s. These thresholds are defined on the difference in closing price of the cryptocurrency as forecasted for the next hour by the trained LSTM model and the actual closing price in the current hour. The chosen values of these hyperparameters are key factors deciding the success of algorithmic trading model. Hence, a key part of the work is to explore various values of these threshold parameters. To get reliable results, we have done five independent model training and got the sum of trades for range of hyperparameters. We present the experimental results in the form of comprehensible data visualizations. The first visual specifically explores the impact of buy T_b on the average cash flow from trades - the results of trades are plotted in Figure 5.21 . A similar plot for determining the average trade cash flow for different values of short-selling thresholds is presented in Figure 5.22. Figure 5.23 presents the combined impact of buy and short selling thresholds on the overall trade positions. This is a scatter plot of all the buy and short selling thresholds, where the colour legend of bubbles is indicative of the magnitude of profit made from all trades on the test set. On conducting the first set of discrete parameter search, we found that the buy threshold of 150 USD and short-selling threshold of -140 USD give the maximum profit for the test set. The cumulative trading cashflow based on algorithmic trades for these optimal thresholds is plotted in Figure 5.24.

To pursue further improvements in the trading profits, we did a more fine-grained search of the threshold hyperparameter T_b in range $[120, 240]$ by increments of 1. Similarly, we did a fine-grained search on the hyperparameter T_s by varying it in the range of $[-150, -100]$. This fine-grained hyperparameter search further results in better trades on test data. This is evident from the fact that from the initial wider hyperparameter search, the best trading position (of all runs) resulted in a profit of USD 20,607.72, while from the second, more fine-grained search the most profitable trade was USD 21,436.42. The best set of hyperparameters from the fine-grained search was T_b = 152 USD, T_s = -115 USD. The plot showing the cumulative sum of trades for this set of hyperparameters is shown in Figure 5.25.

The overall effectiveness of this strategy is demonstrated by the fact that taking the average of all the threshold combinations, over the set of five independent experiment runs, always resulted in a net positive trade for both the initial wider search and the subsequent fine-grained parameter search.

5.4.3 Trading Strategy Hyperparameter Tuning and Results for Ethereum

A good algorithmic trading strategy should not be tied to just one specific financial asset. Hence, as an additional test for the proposed algorithmic trading strategy, we applied it to trades in *Ethereum*, the second most valuable cryptocurrency. We obtained the hourly closing prices for Ethereum for the similar period of study. The

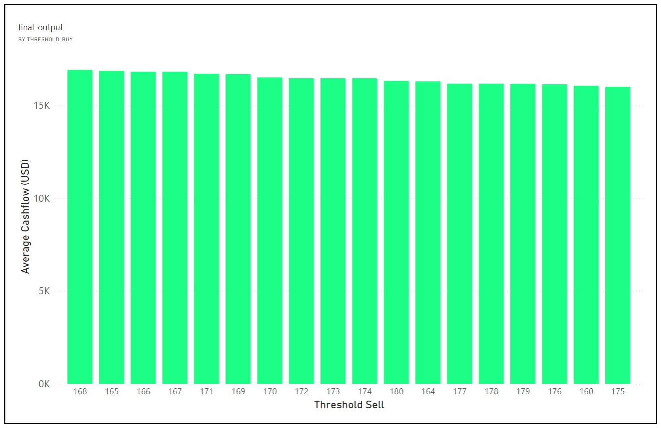

Figure 5.21: A plot showing the average trading cashflow for different values of buy threshold. Note that the average payoff is calculated for five runs with every value of short-sell threshold considered, i.e., (-500, -50] incremented by 10 (Only the threshold values where profits exceed 16 thousand USD are included for clarity)

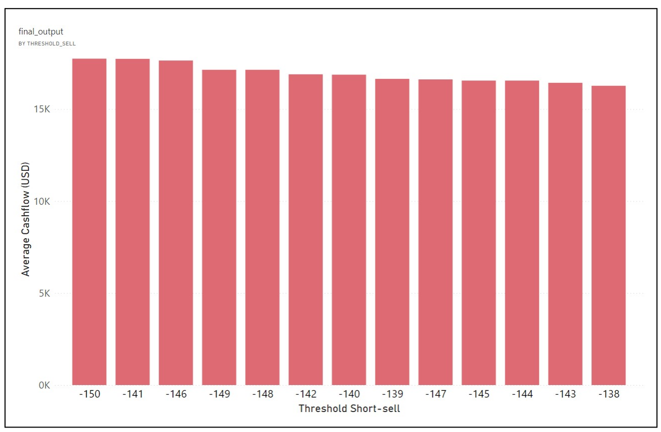

Figure 5.22: A plot showing the average trading cashflow for different values of short selling threshold. Note that the average payoff is calculated for five runs with every value of buy threshold considered, i.e., [50, 500) incremented by 10 (Only the threshold values where profits exceed 16 thousand USD are included for clarity)

strategy also performs equally well for Ethereum, as characterized by the plot of actual vs predicted prices in Figure 5.26 as well as the net positive profits for different set of hyperparameters.

Due to the difference in price scale of Bitcoin and Ethereum (as evident from the Ethereum prices in Figure 5.27), the threshold search for Ethereum was done on

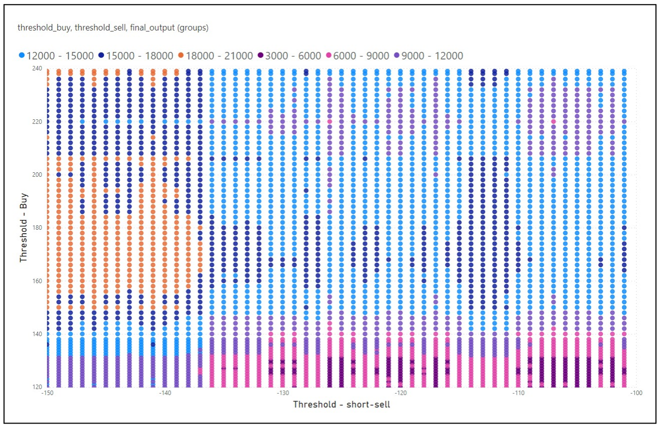

Figure 5.23: A scatter plot showing the average return for different values of buy and short-sell thresholds. The color legend of bubbles is indicative of the magnitude of profit made from all trades on the test set. The orange color of bubble represent the highet profit making thresholds, i.e. between 18,000 USD to 21,000 USD.

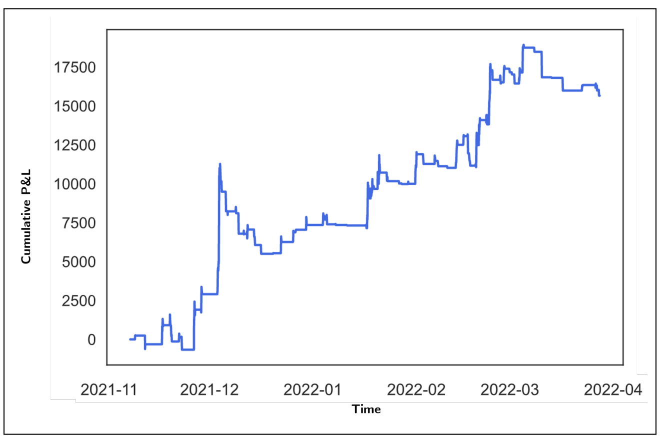

Figure 5.24: Cumulative sum of profit made (in USD) by executing the proposed algorithmic trading strategy for the optimal thresholds found in initial wider search. $T_b = 150$ USD and $T_s = -140$ USD

relatively smaller values. The search space for T_b was [10 USD, 100 USD] with increments of 2, while for T_s, the search space was [-100 USD, -10 USD] with increments of 1. The average trading cashflow for the different values of T_b is plotted in Figure 5.28, while a similar plot for T_s is provided in Figure 5.29. The plots clearly show that the most profit generating values for T_b and T_s are clustered, and have minimal

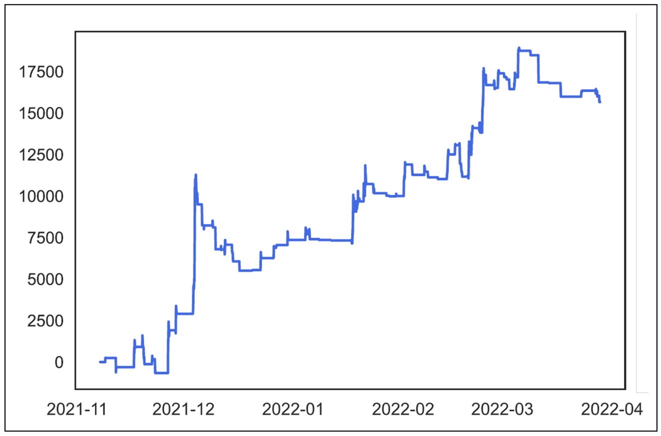

Figure 5.25: Cumulative sum of profit made (in USD) by executing the proposed algorithmic trading strategy for the optimal thresholds found in fine-grained search. $T_b = 152$ USD and $T_s = -115$ USD

relative difference. The cluster is also depicted by the orange colored points in the scatter plot (Figure 5.23). The optimal values of hyperparameters for Ethereum were T_b = 12 USD, and T_s = -11 USD. The cumulative sum of trades for Ethereum for these optimal set of hyperparameters is plotted in Figure 5.30. We see a net positive (P&L), with majority of the movements generating a profit (although there are some loss making trades), which demonstrate the effectiveness of the proposed strategy.

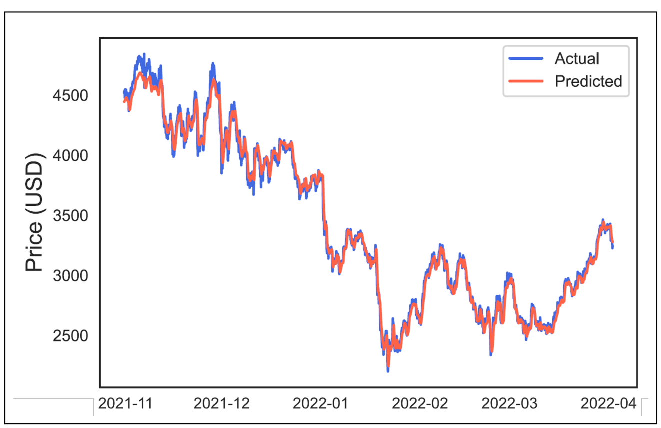

Figure 5.26: Actual vs Forecast prices for test dataset on Ethereum. Predicted values closely follow the Actuals.

Figure 5.27: Hourly Ethereum price data taken from yahoo! finance from the period of 2020-04-03 02:00:00 to 2022-04-01 21:00:00.

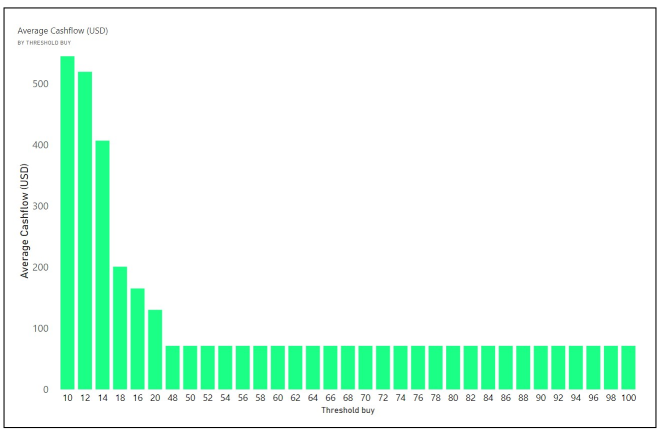

Figure 5.28: A plot showing the average trading cashflow for different values of buy threshold T_b for Ethereum

5.4.4 Extending Strategy to High-Frequency Minute Data

As a brief extension to the work, we also applied the proposed LSTM based algorithmic trading strategy to the high-frequency data available at minute level granularity. The minute-by-minute market data from Yahoo! finance is only available for the preceding 1 week from the time of extraction. The period of data we used for this extended study is from 2022-04-19 05:01:00+00:00 to 2022-04-25 17:42:00+00:00. The initial 80% of this was used for training and the recent 20% was used for testing.

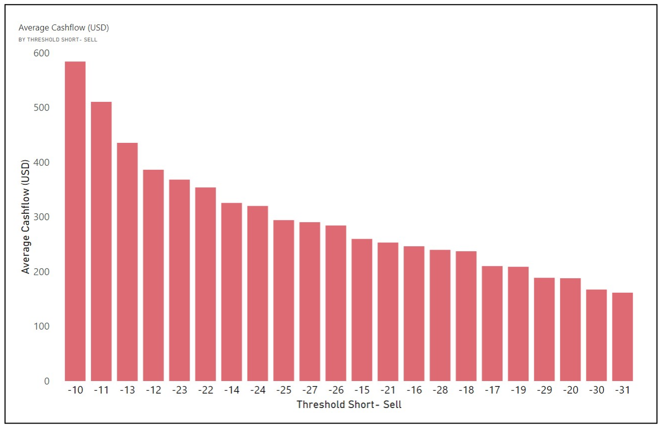

Figure 5.29: A plot showing the average trading cashflow for different values of short selling threshold T_s for Ethereum

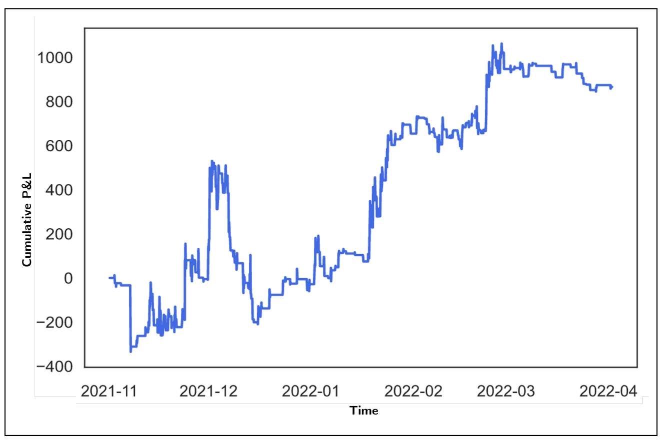

Figure 5.30: Cumulative sum of trades for Ethereum with T_b = 12 USD, and T_s = -11 USD

As can be seen in Figure 5.31, the model forecasts are apt in capturing the trend in actuals for test data. The cumulative profit/ loss (P&L) on high-frequency trading is shown in Figure 5.32.

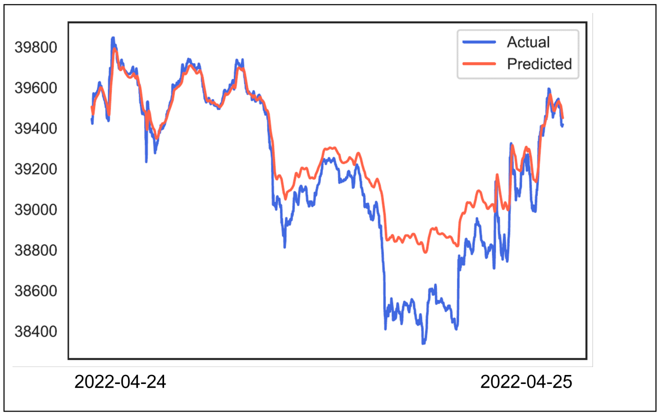

Figure 5.31: Actual vs forecasts on test data for high-frequency minute trading. As can be seen in figure, the model forecasts are apt in capturing the trend in actuals for test data at minute frequency

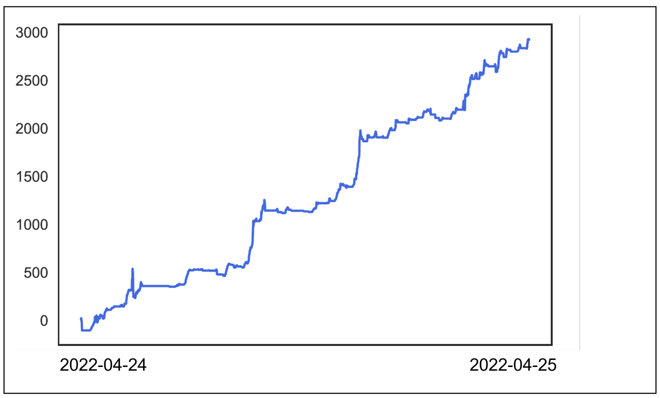

Figure 5.32: Cumulative sum of trades for Bitcoin in high frequency minute data with $T_b = 5$ USD, and $T_s =$ -5 USD. It can be seen that the strategy generates profits which grow over time on the test data.

5.4.5 Strategy Performance with ARIMA as the Forecasting Model

As a further extension of this study, we sought to compare the performance of the LSTM model used in this algorithmic trading strategy with traditional forecasting models. We chose to implement this algorithmic trading strategy with autoregressive integrated moving average (ARIMA) as the underlying forecasting model. We used

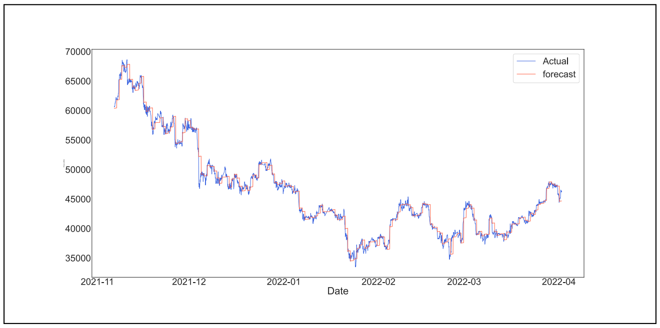

Figure 5.33: Actual vs forecasts for the optimal ARIMA(4, 1, 2) model. These parameters are found by optimizing the model on Akaike Information Criterion (AIC). The model does a fair job as predicted values follow the actuals closely.

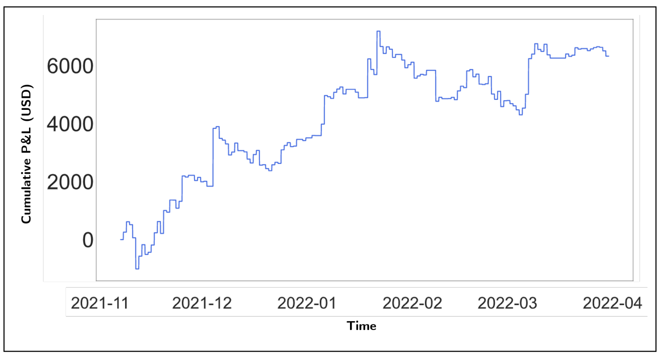

Figure 5.34: The cumulative sum of trades by using the optimal ARIMA(4, 1, 2) for ($T_b = 173, T_s = -102$). The average profits generated with ARIMA model (USD 6,326) are significantly lower than the corresponding profits generated when the LSTM model (USD 21,239)

the hourly dataset for Bitcoin prices, for similar periods as used in the above study for LSTM.

For implementation and tuning of ARIMA model we used the *pmdarima* package available in python. We performed an augmented Dickey-Fuller test (ADF) as well as Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test to determine the differencing term d for the ARIMA model which came out to be 1. We used the *autoarima* function available within the pmdarima package to determine the optimal p and q for ARIMA model. The function found optimal $p = 4$ and $q = 2$. In order to make an equivalent comparison with the LSTM forecasting model, we choose the ARIMA model to make 20 steps ahead forecast (similar to the window size of 20 in LSTM), and hence the ARIMA model updates after every 20 observations. The plot of actual vs forecast price given by the optimal ARIMA model is presented in Figure 5.33. The steps for the algorithmic trading strategy remain similar to the ones mentioned in Chapter 4.3 after the price forecasts are produced by the ARIMA model. The optimal threshold (T_b, T_s) was captured as $(173, -102)$ for the ARIMA model. The cumulative returns produced in this case are depicted in Figure 5.34. It can be clearly seen that the profits are significantly lower than the profits generated when the LSTM model was used as the forecasting core of the strategy. The best thresholds gave a cumulative trading output of USD 21,239, while the similar strategy, on the similar price dataset, using the optimal ARIMA model gave the best profit of 6,326 USD only.

Chapter 6

Decentralized Finance - Review and Study in Risk Analysis

6.1 Review of Decentralized Finance (DeFi)

6.1.1 Overview

The Bitcoin whitepaper[4] published in 2008 proposed a novel decentralized ledger, later called blockchain, which enabled multiple transacting parties to agree upon the shared state of the ledger without a trusted intermediary. Blockchain technology has been used to implement many decentralized payment systems, with the general term Cryptocurrency coined for the native unit of values. The launch of the Turing-complete Ethereum blockchain[58] in 2015 extended the scope of blockchain-based financial systems beyond cryptocurrencies. The suite of non-custodial financial solutions deployed as Smart Contracts over Turing-complete blockchains is broadly called

Decentralized Finance (DeFi). These solutions have gained widespread popularity as investment vehicles in the last two years, with their total value locked (TVL) exceeding USD 100 Billion. This chapter reviews the key financial services offered in DeFi and draws a parallel to the corresponding services in the centralized financial industry. Some technical and economic risks associated with DeFi investments are also discussed in this chapter of the thesis. Most of the existing reviews on DeFi focus on some specific DeFi services, are theoretically inclined, and are intended for academics in computer science or economics. This chapter, on the other hand, aims to give an overview of the current state of the DeFi ecosystem. We aim to keep this review lucid to make it accessible to a broader audience without compromising academic rigor. The intended audience for this chapter includes anyone with a basic understanding of financial markets and blockchain systems. This work will be specifically helpful for investment professionals to understand the rapidly evolving ecosystem of DeFi services.

6.1.2 Motivation

Capital investment is a pillar of the modern economic system. Individuals tend to invest their savings as a means to hedge against inflation. There is a broad ecosystem of organizations that manage the investment from numerous individual investors, pool it, and allocate it across various assets throughout the global financial markets. The economy which receives these investments benefits due to the growth and development it brings along, while for the investors, it leads to the growth of their capital and wealth. Many of these investments involve financial instruments like stocks, bonds, and derivatives. These instruments are also traded independently on global financial exchanges, with their prices varying. The world of finance is primarily digitized, with all the information being stored in digital format. In many cases, these assets are traded with very high frequency by large institutions like pension funds that manage their clients' wealth and aim to give them good returns through their trades. Despite sophisticated risk measures and hedging strategies, these investment institutions may sometimes incur hefty losses (thus affecting their clients). It is especially true at times of extreme events like the 2008 financial crisis and the 2019 pandemic, partly due to the ill effects of a centralized financial system that lacks transparency. Also, economic growth may flatten or decline when capital allocation strategies, controlled by a handful of executives in large institutions, are planned poorly. A new class of financial assets called cryptocurrency was envisaged in 2008 with the launch of the Bitcoin white paper [4] by a person or a group under the pseudonym Satoshi Nakamoto. The subsequent launch of the Bitcoin peer-to-peer network in a decentralized manner. The key achievement of this white paper was the solution to maintain a distributed ledger of transactions among a set of participants and ensure consensus on the ledgers' state without involving a trusted central party. Blockchain technology is based on well-established cryptographic primitives of hashing and public-key encryption. Another major step in blockchain-powered finance was the launch of the Ethereum blockchain network [59]. Ethereum took the core ideas from Bitcoin and extended these to create a general-purpose platform (not just a currency). Ethereum is a Turing-complete blockchain supporting smart contracts that can be programmed using Solidity [60]. Ethereum Virtual Machine (EVM) uses the consensus mechanism of blockchain to maintain a globally coherent state among its participating nodes. The consensus mechanism of blockchain can be seen as a public append-only data structure with the following main properties.

1. Persistence: data cannot be altered once written to the blockchain[1].
2. Consensus: All honest participants have the same data[2].
3. Liveliness: All participants can add new transactions.
4. Openness: Any participant can add data to the blockchain.

Smart contracts, in their basic form, are programs in which a set of encoding rules are enforced by a blockchain's consensus mechanism(s). The distributed framework allows trustless economic interactions between parties. The Ethereum blockchain embedded the first working implementation of smart contracts. Following Ethereum, other blockchains such as Binance (BNB) [61], Cardano (ADA) [62], Solana (SOL) [63] and Avalanche (AVAX) [64] with smart-contract capabilities provided other platforms to build decentralized applications using an underlying blockchain as core consensus layer. The concept of decentralized autonomous organizations (DAOs) was subsequently developed along this line. DAOs are companies that are governed by their token holders and use the blockchain to manage token ownership. Another significant development was the introduction of decentralized finance (DeFi) solutions, which involve building a complete financial services ecosystem (mirroring the centralized version, which includes some core institutions like banks and exchanges) over

[1]Assuming honest nodes controls more than 50% of the network
[2]They might diverge for recently added blocks, but the consensus is guaranteed for older blocks

blockchains based on smart contracts. In the past two years, the interest in DeFi has exploded, with total value locked (TVL) reaching more than USD 100 Billion [65]. A key advantage of DeFi over centralized institutions is that all transactions are public and posted on the underlying blockchain[3]. This makes the underlying smart contracts very transparent and auditable. Moreover, to attract investment, the DeFi protocols may have specific incentives to reward the initial investors, making investing in the underlying protocol more appealing. Even though the value locked in DeFi is a tiny portion of the centralized financial institutions, it has the potential to take a significant share of the market. Due to its decentralized, globally accessible 24/7/365, openly auditable nature, and non-custodial architecture that can offer new financial products, DeFi has the potential to resolve the existing inefficiencies of capital allocation in today's centralized financial ecosystem. The volume of investment in DeFi is growing exponentially, but unlike traditional finance, various statistical measures to quantify investment volatility and risk exposure have not yet been devised. Although the data in the blockchain are public and universally accessible, it is still in a raw format, which needs to be aggregated and extrapolated to provide helpful information and support investment decisions[66]. In this chapter, we draw a big picture of DeFi's state of the art and bridge the gap between traditional financial services and DeFi applications.

[3]This is true for all of the mentioned Blockchains except Solana, wherein the historical transactions are purged after a certain period, while only the roll-ups kept. Also, the layer2 blockchain networks sometimes do the same with Ethereum

6.1.3 Brief Overview of Traditional Finance Ecosystem

We will review some of the critical entities of the financial system and the essential services they provide for running today's market-based economies. The key characteristic of these systems is centralized control and the requirement of a trusted intermediary to make financial transactions. The key entities in the traditional financial system and their interconnections are listed below.

- Central Banks: The institution responsible for deciding the overall monetary policy of an economy, from monetary supply to interest rates. Typically, a central bank also supervises commercial banks and non-banking public financial institutions. Examples of central banks are the Bank of England in the UK [67] and the Federal Reserve Board in the USA [68].

- Financial Regulators: These are the authorities typically controlled by the government to oversee the financial activities within their jurisdiction. Different regulators can exist within a jurisdiction, each dedicated to monitoring specific economic activities. Their stated goal is to ensure fairness and prevent fraudulent activities. For example, the Financial Conduct Authority (FCA) [69] is the financial services regulator in the UK and the Securities and Exchange Commission (SEC) [70] is the securities regulator in the USA.

- Exchanges: An exchange is a platform for trading financial instruments like stocks, derivatives, bonds, etc. It also acts as a medium for companies to raise capital from investors by getting listed. The governing body of exchanges also must ensure a fair marketplace for investors. Almost all modern-day exchanges are functioning electronically.

- Commercial Banks: These are the institutions that provide the banking services (like savings, borrowing, etc.) to individuals, companies, and organizations. They make a bridge between the clients having excess capital by accepting deposits through saving services and the clients who need money by offering them lending services.

- Brokers: These individuals or companies act as intermediaries between investors and financial exchanges. They facilitate the individuals to trade assets in exchange for commission/ fees.

- Asset Management Companies: These entities manage their clients' pooled funds. These funds include hedge funds, mutual funds, pension funds, private funds from High Net worth Individuals (HNIs), etc.

Financial services are essential pillars of a modern economy. These services ensure that the global economic system can run and grow, and hence people can improve their living standards by participating in this system. The form of the financial system varies between classes of participants. Generally, individuals aim to build wealth and improve their living standards, while businesses strive to get resources to invest in productive activities while earning profit from these activities. An ideal financial system ensures efficient resource allocation across the participating actors to achieve these goals. The distribution efficiency among participants is measured in terms of utility derived by each of them. The key components of the financial system are financial contracts that determine how real resources will be allocated among participants. The legal system that enforces these contracts and the regulator that oversees the entire system detect and rectify irregularities by enforcing investor rights and preventing bad actors from using the system by enforcing Know-Your-Customer (KYC) and Anti-Money-Laundering (AML) standards. Also, the financial system performs risk management by mandating specific minimum capital requirements for institutions. Some of the factors that can undermine a financial system are [71]:

1. Lack of trading opportunities: Inefficacy of the financial system reduces the composability of trading strategies that could increase the utility of counterparties.

2. High systemic risk: a systemic risk is realized when multiple participating entities may collapse one after another, like a domino, due to high interdependence.

3. Inefficient split in the trade benefits resulting in monopolies.

On the contrary, some of the factors that contribute to a healthy financial system are [71]:

1. Allocation: An ideal financial system has resource distribution that optimizes the increase in utility of participants

2. Inclusiveness: Actors willingly participate in an ideal financial system, and the system provides enough opportunities for participation of new actors.

3. Unbiased regulation: Regulators must ensure that the system's spillovers are managed in everyone's best interest.

6.1.4 Bitcoin and Decentralized Currency Systems

The standard centralized payment method through government-issued and controlled banknotes has been a norm since it replaced the barter trade system. This norm was questioned by Satoshi Nakamoto (2008) in his seminal paper [4] that advocated for a decentralized system of payment over the traditional intermediate trusted party system (e.g., a central bank). Nakamoto highlighted certain drawbacks of the conventional online transaction systems, such as cost incurred for transactions, minimum transaction limit, and the fact that transactions can be reversible. Subsequently, they proposed a decentralized digital currency, called Bitcoin, that uses cryptography to carry out transactions and securely handle ownership. These transactions are stored in a decentralized system called a blockchain. A blockchain is a list of blocks where the miners record different transactions as blocks, and each block contains a list of validated transactions (after being added to the blockchain). The miners are rewarded with some bitcoins and obtain transaction fees from users once they add a block to the blockchain. Multiple blogs, opinion columns, and articles are published in various conferences and journals ranging from Business to Law and Computer Science to Finance, which discuss Cryptocurrencies at different levels and facets. This chapter of the thesis does not attempt a thorough review of all aspects (business, computing, finance, and law) of cryptocurrencies. We refer to Chohan [72] for a brief thematic review of cryptocurrency markets. Bitcoin [4] was one of the first payment methods based on a peer-to-peer network that allows transfers without a traditional trusted third party like a central bank or other government institution. It was followed by Ethereum [59] as the first blockchain network with smart-contract capabilities. Commonly used payment systems are based on a trust-based model where a financial institution is responsible for mediating all transactions. Blockchain networks use cryptography algorithms that validate transactions and prevent double-spending transactions using an underlying consensus algorithm [73; 74]. Once a block is inserted into the chain, it cannot be changed unless the majority of the participants agree according to the underlying consensus protocol. This is because, as more and more blocks are inserted into the chain, the amount of network coordination to change becomes a daunting task. Thus, blockchain systems are reliable as long as honest nodes constitute a majority of the network. To guarantee that a majority of the network is honest, an incentive is given for each validated block inserted into the blockchain.

6.1.5 Smart Contracts

Nick Szabo introduced the concept of the smart contract[75], and suggested that the terms of a legal contract could be embedded in code, which would execute it autonomously without the requirement of a third party. Szabo used as an example a vending machine, which can handle a simple logic such as "input $\Rightarrow$ selection $\Rightarrow$ authorization $\Rightarrow$ change" without any human intermediary. With the introduction of blockchain technology, smart contracts became popular as complex programs deployed on transaction-based blockchains. They consist of rules verifying, controlling, and self-executing a predefined agreement. As they are executed on a decentralized blockchain network that is transparent, traceable, and irreversible, smart contracts often involve anonymous parties in a trustless setting without the participation of third parties. Smart contracts allow a deterministic, rapid, and cost-efficient execution of contracts between parties. A smart contract has an address used to call the program, functions which encode its behavior, and data that maintains the state among all the nodes in the network. To execute a smart contract, the users must pay a fee, usually identified in terms of "gas". The amount of gas varies depending on the smart contract's complexity and cost for each operation it executes. Gas is generally paid with the underlying currency of the blockchain on which the smart contract is running. In the case of Ethereum, the denomination is Gwei, while the currency is Ether [76].

6.1.6 ERC-20 and ERC-721 Tokens

Smart contracts enable developers to implement tokens, and a variety of complex distributed programs such as lending platforms, e.g. Aave [77], Compound [78], and decentralized exchanges (DEXes), such as Uniswap [79], SushiSwap [80]. Tokens are a standard implementation of smart contracts. The most common tokens constitute ERC-20 and ERC-721 standards for fungible and non-fungible tokens on the Ethereum blockchain. A token can be fungible or non-fungible. Fungible tokens are interchangeable as all the tokens in circulation have the same value, while the non-fungible token (NFT) is unique, and each has its own value [81].

6.1.7 Key DeFi Services

6.1.7.1 Borrowing/Lending

Lending is a vital component of economic machinery. One of the primary mechanisms is to facilitate the capital exchange agreement between parties with excess capital (called lenders) and parties that need money (called borrowers). Lending is a mutually beneficial agreement where the borrowed capital must be returned to the lender along with additional payment in the form of interest as per an agreed-upon timeline [82]. Also, to hedge against the risk of non-repayment of the lent amount, the lender takes custody of some asset called *collateral* from the borrower, which the lender can monetize to cover the unpaid loaned amount. Decentralized lending is a construct like traditional lending for digital assets like cryptocurrencies and tokens hosted on blockchain platforms. But, unlike lending in conventional finance, the decentralized lending platforms cannot accept off-chain assets as collateral [83]. Moreover, as blockchain-based digital assets are more volatile when compared to most traditional financial assets, their lending has to be over-collateralized, i.e., the value of the collateral token at the time of the lending has to be greater than the lent token. The over-collateralized agreements can be seen in the issue of DAI stablecoins by MakerDAO [84] as well as popular decentralized lending platforms like Aave [77] and Compound [78]. Under-collateralized lending protocols such as Alpha Homora [85] also exist, but in such platforms, there are many restrictions on spending of the borrowed funds, and the ownership of the funds stays with the lending pool instead of being transferred to the borrower. Liquidation is an exciting mechanism associated with decentralized lending, which allows a third party to buy the collateral from the lending pool at a discounted price in case the value of the collateral falls below a certain threshold relative to the borrowed asset [83]. It acts as a risk management mechanism for the borrower while providing liquidators with profit-making opportunities. Flash loans are a novel risk-free lending mechanism introduced in DeFi, which can only be implemented in blockchain-based settlement systems and is not available in traditional finance. A flash loan involves lending a digital asset and its subsequent repayment within a single atomic transaction [86]. If the borrowed amount is not repaid, the entire transaction is reverted and not included in the block, which is equivalent to the loan event not taking place. One of the most common applications of flash loans is to gather funds for utilizing the arbitrage opportunities between different Decentralized Exchanges to earn risk-free profit [87].

6.1.7.2 Stablecoins

One of the key properties of money in modern economic systems is that it is a "store of value" [88]. A commodity or asset can be considered a value store if it can be reliably saved, retrieved, and exchanged in future times while also being predictably useful as a medium of exchange on retrieval [89]. In other words, its value should remain considerably stable with time. But cryptocurrencies, which are the native medium of exchange on their respective blockchains, are too volatile to be considered a reliable store of value. As the government-backed fiat currencies act as means of payment for all day-to-day financial transactions in the real economy, an equivalent, less volatile store of value is needed for the transactions on the blockchain-powered financial ecosystem. Stablecoins [90] are designed to fill this gap. These are the smart contract-based digital tokens deployed over the blockchain whose value is pegged to the non-volatile assets like fiat currencies. Hence, they act as stable value stores for payments settled on blockchains. Most of the widely used stablecoins are pegged to the US dollar, the most circulated fiat currency in the world. The first stablecoin was Tether, pegged to USD and launched in 2015 [91]. Stablecoins can be custodial or non-custodial. Custodial stablecoins rely on a trusting third party to maintain the stability in prices, generally by off-chain collateral backing of the underlying asset like US Dollars. Non-custodial stablecoins, on the other hand, use economic mechanisms to maintain the peg to stable assets [92]. Tether, USDC, and Binance USD are examples of custodial stablecoins, while MakerDAO is a leading example of non-custodial stablecoins [91].

6.1.7.3 Decentralized Exchanges (DEXes)

A financial exchange is a platform on which financial assets, either traditional or blockchain-based digital assets, are traded by different parties. Participating traders enter their quotes to buy or sell a particular asset in the traditional exchange model. The exchange system has two main steps: first, to match the trades which can be executed, and second, the settlement between counterparties. Such an exchange is called an order book and requires a central authority to accept and match quotes and act as an escrow for the financial assets of the counterparties until the trade is executed. The centralized cryptocurrency/token exchanges like Coinbase, Binance, and FTX are also based on an order book model where trade matching and settlement are carried out on the centralized server of the platform service provider. An order book based decentralized exchange can also be set up in the form of smart contracts deployed over the blockchain. EtherDelta is an example of such an approach but suffers from various problems like latency, high gas fees, miner front running, etc. [93], even though it provides certain advantages of decentralized finance in the form of censorship resistance and robustness. Another class of exchange models more suitable for blockchain-based decentralized financial assets is the Automated Market Makers (AMMs) [94]. The most popular of them is the Constant Function Market Makers (CFMM) [95], which maintain a mathematical invariant (for example, a product of the quantity of assets) during the trade. Unlike order book exchange, in CFMM, transactions happen between a trader and a pool of funds being traded, a smart contract, rather than directly between trading parties. A separate class of investors that provides liquidity in the pool is called Liquidity Providers (LP), which is incentivized by awarding fees accrued on trades in proportion to the ownership of pool reserves. The CFMMs are relatively simple to implement as smart contracts and incur less gas fees than the order book based models. Unfortunately, CFFMs have significant drawbacks when compared to the order book exchange, such as high slippage, impermanent loss, and miner manipulations [93]. However, the trade-off for DEX implementation is mainly in favor of CFMMs.

6.1.8 Investment Opportunities in DeFi Ecosystem

The DeFi sector is still evolving and is a high-risk and high-return investment ecosystem. There exist several investment opportunities in this alternative financial system. The major ones are described below.

6.1.8.1 Liquidity Provider (LP)

A liquidity provider (LP) is a key agent that enables the functioning of Automated Market Maker (AMM) based Decentralized Exchanges (DEX) across different blockchain platforms by providing liquidity in the form of digital tokens. A trader, another essential factor, uses the pooled liquidity to trade one pooled token with another. As an incentive to the liquidity provider, most protocols give a fixed fee specified as a percentage of the token being swapped on the platform. In addition to the fixed fee, in some pools, the liquidity providers receive additional reward tokens by either the DEX protocol governance or by the governance of one of the token contracts in the pool to attract liquidity.

6.1.8.2 Arbitrage

In general economic terms, arbitrage is the process of taking advantage of the price difference of an asset in different markets by buying at a lower cost from one market and selling at a higher price to another market, thus making a risk-free profit until the prices in both markets become equal, which is generally a short period. The process of arbitrage ensures the equilibrium of asset prices across all markets. Traditional markets are usually very efficient and with limited arbitrage opportunities, but the DeFi markets are still developing and provide plenty of arbitrage opportunities. The primary modus operandi of executing arbitrage is exploiting a token's price difference across different DEX platforms, which may exist due to market inefficiencies. The decentralized lending platforms' flash loan service can provide collateral-free capital required to book a profit through DEX arbitrage, which involves buying a token at a lower price from one exchange and selling it at a higher price to another exchange.

6.1.8.3 Liquidation Bots

As discussed in Section 6.1.7.1 (Decentralized Borrowing/ Lending), the leading platforms like Compound and Aave have a protective mechanism called *liquidation* in place, which prevents the risk of collateral depreciating below the lent amount. When the collateral value falls below a certain predefined threshold, this mechanism allows the lending contract to sell it to any willing buyer at a discounted price to incentivize the buyer. This discount becomes an investment opportunity for the *liquidators*, as they can buy a token at lower than market prices. As per an analysis conducted by Gudgeon et al. in 2020 [83] on the *Compound* lending protocol, liquidators have become very efficient over time, with over 70 percent of liquidable positions getting immediately liquidated. This efficiency is possible due to specialized computer programs called liquidation bots, which keep parsing the state of the blockchain to look for potentially profitable liquidation opportunities and execute a liquidation transaction. The empirical study by Qin et al.[96] demonstrates algorithmic strategies, which the liquidation bots can use to make profitable liquidations across some popular decentralized lending platforms.

6.1.8.4 HODL

HODL stands for 'Hold On For Dear Life', which refers to people that have a long-term investment strategy in cryptocurrencies and will hold their tokens regardless of market volatility or downturn. The investor following this strategy is called a *HOLDer*. The strong belief of HODLer is that despite the short-term volatility in price movements, the long-term trend is that price of the asset would go up. This is analogous to *value investing* in traditional finance. To compare investment strategies, one can contrast between HODLing and Liquidity provision. While a HODLer earns returns by the value appreciation in the digital asset in their custody, a liquidity provider holding the same asset can stake it in some liquidity pool and earn fees from the trades executed over the pool. Providing liquidity, however, entails an additional risk of *impermanent loss*, which is briefly explained in the next section.

6.1.9 Risks Specific to DeFi

As mentioned earlier, DeFi is a high-risk and high-return ecosystem. In addition to the financial risks involved with traditional assets, trading in DeFi assets entails additional risks native to the ecosystem, as described below.

6.1.9.1 Bugs/Hacks

The Decentralized Finance protocols are implemented as smart contracts deployed over a Turing-complete blockchain platform. Like any other piece of code, smart contracts are vulnerable to potential bugs and hacks, and hackers can exploit a bug in a smart contract to drain investors' funds from the contract. The most famous example of a smart contract bug being exploited was *The DAO Hack,* which drained around USD 60 Million worth of Ether from *The DAO* smart contract on the Ethereum blockchain within the month of its launch in 2016. The hackers exploited the *re-entrancy vulnerability*, a design flaw in the smart contract that allowed them to recursively withdraw funds from the smart contract without updating their remaining balance. Different technical bugs might exist in smart contracts and can put the investors' funds in danger of exploitation by malicious actors. Hence, the DeFi protocols hire independent contract auditors to check for vulnerabilities and make the finding of these audits public to increase investor confidence.

6.1.9.2 Miner Extractable Value (MEV)

Miner Extractable Value (MEV) attacks are a risk class associated with transaction order within a block of the underlying blockchain. The risk is based on the premise that ordering transactions within a block can impact the trade returns for the transacting entity. The miner decides the transaction order, and the transactions offering higher gas prices get included on priority in the block to be mined. There are actors called *searchers* which continuously parse the *mempool* of pending transactions to find profit-making opportunities by exploiting MEV. If a profit-making transaction is found, the searcher bots create a new transaction using the identical profit-making strategy and send it to miners with more gas prices than the original transaction. This result in searchers' duplicate transaction getting mined instead of the original transaction, and the profit which was supposed to be earned by the creator of the original transaction is instead taken by the searcher. An even more serious MEV-based attack is *sandwich attack*, in which the searcher earns profit by exploiting the *slippage* in AMM-based DEX at the cost of losses incurred to the originator of the searched transaction.

6.1.9.3 Impermanent Loss

The Liquidity Pools (LPs) bear the risk of impermanent loss in AMM-based Decentralized Exchanges. The LPs can incur a loss due to a change in the reserve ratio of the pool resulting from a divergence in the unit price of the pool tokens. Mathematically, *impermanent loss* is the difference between the current market value of tokens that the LP initially staked in a pool and the current value of pool assets owned by the LP. The impermanent loss incurred by LP can be attributed to the profit of arbitragers who trade out the token with an increasing price in return for the token, which is losing relative value.

6.1.9.4 Liquidations

The liquidation risk is a risk that borrowers face in the decentralized lending protocols due to the liquidation mechanisms in these protocols. Almost all the lending protocols are over-collateralized to counter the risk of the high volatility of crypto assets. Hence, these protocols require the borrower to provide the collateral tokens with a value greater than the borrowed tokens by some minimum ratio called *liquidation threshold*. Suppose the value of collateral relative to the loan falls below this ratio. In that case, liquidation kicks in, and any third party, called a liquidator, can buy the collateral at a discounted price as described in Section 6.1.8.3. The borrower can either add more tokens as collateral to bring it above the liquidation ratio or prepare to lose the collateral to liquidators.

6.1.9.5 Fraudulent Projects/Tokens

Fraudulent Projects and Tokens constitute one of the most common and gravest risks in the decentralized finance investment space. Being permissionless is one of the four key tenets of DeFi; but it is also the reason to be cautious as an investor. Since anyone with Internet access can start a new token on DeFi ecosystem without any KYC requirements and promote it as valuable on social media platforms, an investor must be able to identify frauds from genuine tokens. As per a CNBC report [97], over 10 Billion USD fraud might have occurred in DeFi space in the year 2021 alone. The most common category of DeFi scam is *rug pull,* which accounted for around 37 percent [98] of all cryptocurrency-based scams in 2021. The general scheme in conducting a rug pull starts with launching a token native to a blockchain platform, like an ERC-20 token on Ethereum. The token is then listed on a Decentralized Exchange, such as Uniswap on the Ethereum blockchain. A few days after listing, the token creator (scammer) starts marketing to generate demand for the token. As the token starts appreciating in value, the token creator adds their scam token to one or more liquidity pools and pairs it with a valuable token such as Ethereum or USDC. The marketing campaigns continue, generating demand and increasing token prices. Scammers even use incentive programs to encourage participation. For example, everyone participating in the liquidity pool gets an extra 1000 tokens per week. Once there is sufficient liquidity in the pool for the scammer to profit, the scammer sells all their scam tokens for valuable tokens, leaving other participants with only scam tokens in the pool. Another variation of the scam rug pull involves using a proxy, or upgradable, smart contracts. A scammer creates a token using a smart contract that executes code that is not immutably written on the blockchain but stored on a proxy server under the scammer's control. They follow the same marketing methodology previously described. Instead of using liquidity pools, they sell scam tokens for fiat currencies or swap them for valuable tokens. Then one day, they alter the off-chain code, and the scam token stops working, and existing holdings become worthless. It is complicated to tell a well-executed scam token apart from a genuine token. Smart contract code, publications of an independent audit report, websites, white papers, social media activity, and the individuals and companies that created the token should be thoroughly researched before investing in any new token.

6.1.9.6 Regulatory Risks

In all major jurisdictions worldwide, financial services come under the purview of a strict regulatory framework. This is done to protect investors from getting dumped by financial institutions. Financial regulations are necessary to create trust in the financial ecosystem and keep it running. DeFi is designed to replicate the traditional financial services over blockchain more inclusively and openly. Due to its pseudonymous identity management, it is challenging to target individuals in DeFi-related wrongdoings. Instead, regulators worldwide are increasingly inclined toward bringing DeFi services under their purview and consider developing a separate framework around DeFi services. This is evident from SEC charging Zachary Coburn, founder of EtherDelta, an order book based decentralized exchange on Ethereum, over investor fraud[99]. The primary liability under these regulations will be targeted at founders, developers and contract developers. Still, the secondary liability may come to users of protocols, thus making them potential targets of litigation. Also, in extreme cases, the regulators can choose to ban the system altogether, such as the Chinese government declaring the use of significant cryptocurrencies and, by extension, all major DeFi services illegal[100].

6.1.10 Concluding the DeFi Review

The introduction of new technologies could reshape the entire financial ecosystem by unleashing competitive threats to the existing players and allowing new entrants to thrive. In the future, Cryptocurrencies will further challenge cash and fiat currencies. Narrow banks that hold only liquid government bonds will challenge traditional banks built on the fractional-reserve model. Distributed payment systems like those that allow individuals to make payments through blockchains will compete with existing centralized systems anchored in physical public and private banks. Decentralized finance that removes intermediaries will test the role of traditional Wall Street giants and their law firms. To conclude, the financial ecosystem is undergoing a frenetic evolution, and when the dust settles, it could be changed beyond recognition.

6.2 Risk Analysis of LP Returns in an AMM DEX - a Case Study

In this brief study, we quantify the market risk for a liquidity provider in DEX pools. The mechanism of liquidity provider's investment has been described in section 6.1.8.1 . We use the commonly used risk metrics Value at Risk (VaR) and Expected

Shortfall (ES) (see Chapter 3.1 for details) to quantify the risk. Furthermore, we used two methods - historical simulation and a parametric approach by fitting T-distribution to estimate the risk metrics.

6.2.1 Background and Motivation

The total value of the investment in the DeFi ecosystem has grown exponentially in past few years [101]. One of the key investment options in this ecosystem is the provisioning of liquidity in an AMM-based DEX such as *Uniswap V3*. Despite the growing transaction activity and value locked in AMM-based DEXes across multiple blockchain platforms, we did not find any past research study on quantifying the market risk of investment returns of Liquidity providers in AMM DEXes. This brief study is just a small step in this direction. We also compare the quantitative risk of providing liquidity to DEX against just holding the corresponding ERC-20 token pair.

6.2.2 Methodology

Due to the complexity and volume of transactions posted on public blockchains such as Ethereum, the task of filtering and parsing the transactions specific to a specific smart contract belonging to the chosen DEX pool is a complex data engineering process. Hence, we used the data available from the FLUIDEFI [66] multi-chain DeFi investment management and data aggregation platform, which has an extensive data pipeline to capture and parse the transactions from blocks as well as computation of investment returns based on the mathematics defined in protocol whitepapers (for

Table 6.1: Description of Uniswap V3 pools chosen for this study. Pool description includes the ERC-20 token pairs, followed by the size of Tick and Swap fee for the pool. The second column is the address of the corresponding pool in Ethereum mainnet

Pool Description	**Smart Contract Address on Ethereum Mainnet**
USDC-WETH Tick:60 Fee:3%	0x8ad599c3A0ff1De082011EFDDc58f1908eb6e6D8
SHIB-WETH Tick:60 Fee:3%	0x2F62f2B4c5fcd7570a709DeC05D68EA19c82A9ec
WBTC-WETH Tick:10 Fee:0.5%	0x4585FE77225b41b697C938B018E2Ac67Ac5a20c0
WETH-USDT Tick:10 Fee:0.5%	0x11b815efB8f581194ae79006d24E0d814B7697F6
WETH-LOOKS Tick:60 Fee:3%	0x4b5Ab61593A2401B1075b90c04cBCDD3F87CE011
WETH-USDT Tick:60 Fee:3%	0x4e68Ccd3E89f51C3074ca5072bbAC773960dFa36
HEX-USDC Tick:60 Fee:3%	0x69D91B94f0AaF8e8A2586909fA77A5c2c89818d5
DAI-WETH Tick:10 Fee:0.5%	0x60594a405d53811d3BC4766596EFD80fd545A270
USDC-WETH Tick:10 Fee:0.5%	0x88e6A0c2dDD26FEEb64F039a2c41296FcB3f5640
HEX-WETH Tick:60 Fee:3%	0x9e0905249CeEFfFB9605E034b534544684A58BE6

instance Uniswap V3 whitepaper [79]). We choose Uniswap V3 [102] deployed on *Ethereum mainnet* as the DEX to analyse because it is the top DEX by transaction volume [103] during the period of our study. Uniswap V3 allows for flexibility in liquidity provision due to concentrated liquidity positions, which may be advantageous for liquidity providers depending on their risk tolerance or trading strategy. The selection of the liquidity pools for this study from a large set of pools deployed in Uniswap V3 is based on the activity level, which is quantified by the sum of the four key events in Uniswap pools, i.e., *mint*, *burn*, *swap*, and *flash* [79]. The list of these top 10 pools used for the study is provided in table 6.1.

6.2.3 Analysis of LP Returns

6.2.3.1 LP Returns Computation

The investment returns for a LP are denominated in the two constituent tokens of the liquidity pool. Hence, it is impossible to mine the returns based just on the transactions of the liquidity pool. We also need price data for the constituent tokens. FLUIDEFI [66] platform has a completely decentralized pricing engine, which relies entirely on on-chain data to calculate token prices. The pricing engine and smart contract transactions are used in conjunction to calculate LP returns.

Another key point to note is that Uniswap V3 protocol allows concentrated liquidity provision, which implies an uneven distribution of liquidity between different price ranges. To get combined returns for all LPs inside a pool, the platform provides an average of returns for the liquidity spread across all price ranges.

For this analysis, we consider two separate forms of returns earned by an LP. **Price returns** is simply the return on investment due to the price movement of the underlying tokens. This amount would be earned if LP just bought the constituent tokens and held them in its wallet, instead of providing them as liquidity in DEX pool. **Fee returns** are earned when a liquidity provider provides liquidity in a DEX pool. For all the trades (swap events) happening in the pool. Each liquidity pool has a fixed swap fee, which is collected from the wallet making a swap and shared among the liquidity providers in proportion to the contribution to the liquidity pool. Fee return is always positive as there cannot be a negative fee. **Gross return** is the combination of price return, fee return and impermanent loss (see section 6.1.9.3 for details on impermanent loss). This is the actual average return earned by the liquidity providers by investing in a pool.

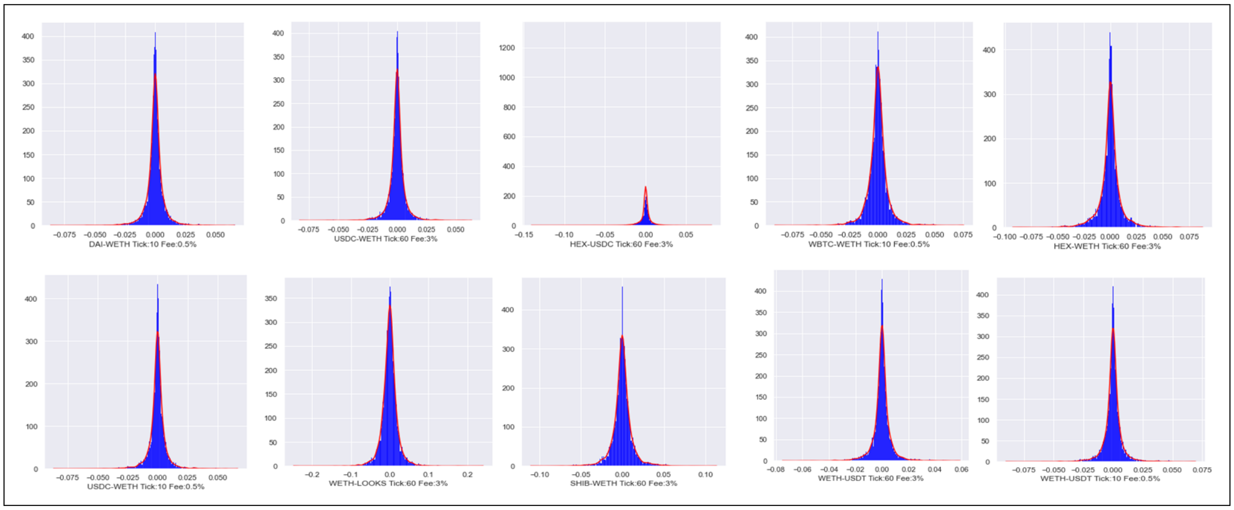

Figure 6.1: Price returns distribution (blue bars) and smoothed curve (red line) of the 10 Uniswap V3 liquidity pools studied. These plots suggest that the price return from the Ethereum ERC-20 tokens follows a similar distribution to that of cryptocurrency returns.

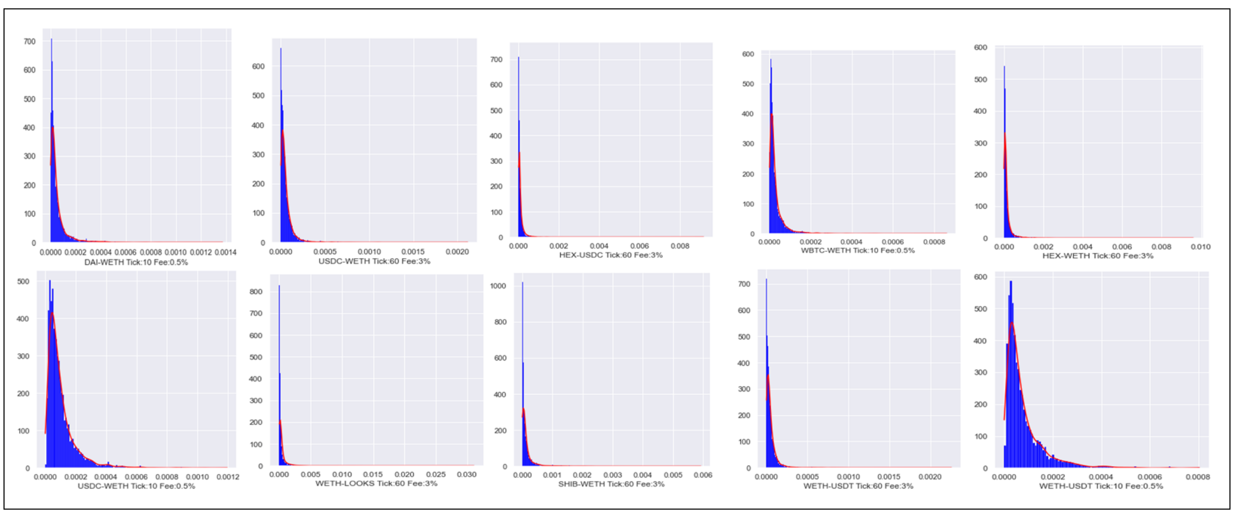

Figure 6.2: Fee returns distribution (blue bars) and smoothed curve (red line) of the 10 Uniswap V3 liquidity pools studied. The fee returns can not be negative hence this distribution is similar to a semi-normal distribution.

6.2.3.2 Observations on LP Returns

We plotted a histogram and fitted a distribution curve on the price returns, fee returns and gross returns for each of the ten liquidity pools studied. The plots for

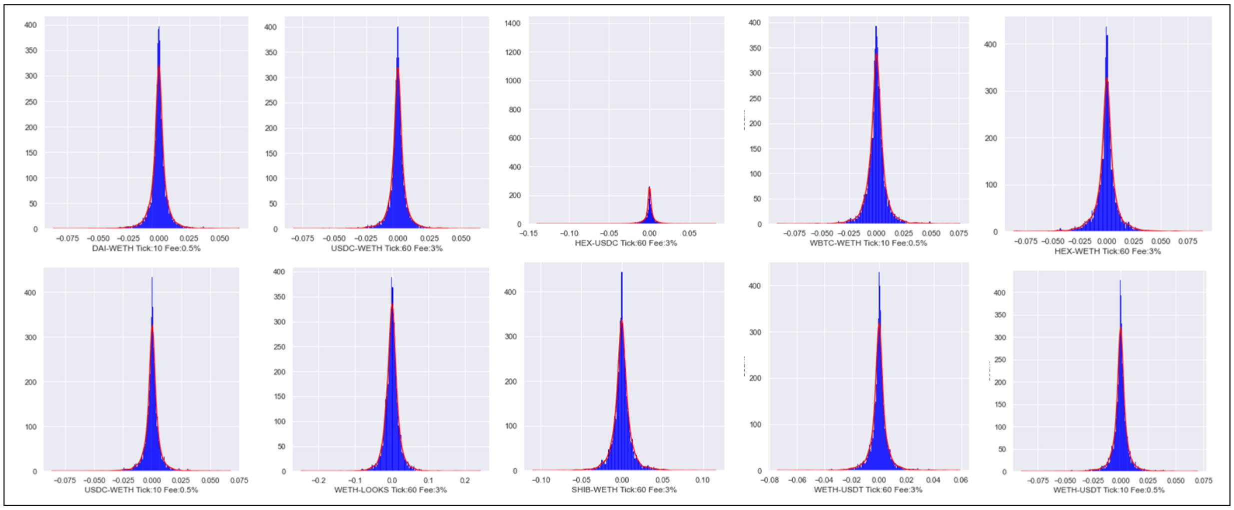

Figure 6.3: Gross returns distribution (blue bars) and smoothed curve (red line) of the 10 Uniswap V3 liquidity pools studied. The gross return distribution is symmetrical around zero. This would imply that the price returns are dominating factor in gross returns.

price returns are shown in 6.1, while that of fee returns in Figure 6.2 and gross returns in Figure 6.3.

It is clear that the price returns fit into a normal distribution centred at zero. This is typical for almost every financial asset like stocks and cryptocurrencies. Also, for some tokens, the distribution is more fat-tailed than others. Specifically, the stablecoin pool (HEX-USDC Tick:60 Fee:3%) has much narrow distribution than all other non-stablecoin pools, which demonstrates the price stability among the stablecoin pair. Moreover, the fee returns, which are always positive, fit into a semi-normal distribution as shown in Figure 6.2. Here also, we can see the different spreads between pools, with the spread being directly proportional to the variation in fee earned, and the area under the curve is proportional to the fee earned. The gross returns for an LP, which is a composite of price returns, fee returns, and impermanent loss are plotted in 6.3 for ten pools being studied. The gross returns also follow normal distribution centred around zero. We will quantify the risk of price returns and compare them to gross returns in the next section.

6.2.3.3 Risk Analysis of LP Returns

We did risk estimation of price returns and gross returns, followed by a comparison between the two. The question we seek to answer is from a quantitative risk perspective, whether providing liquidity to a DEX liquidity pool is a better option than just holding the two underlying tokens.

To answer this question, we compute two widely used quantitative metrics for market risk - Value at Risk (VaR) and Expected Shortfall (ES). The price returns are indicative of the first scenario where the investor just holds the underlying tokens, while gross returns represent the second scenario of providing liquidity to a DEX pool. We use two different models for the risk estimation, the first one is historical simulation (see [104] for details), while the second method is a parametric approach based on fitting a T-distribution on historical returns (see [105] for details). The table 6.2 contains the VaR estimations using historical simulation method, while 6.3 tabulates the ES estimation using the similar method. The tables 6.4 and 6.5 contain the estimations of VaR and ES respectively using the parametric variance-covariance method.

6.2.4 Key Findings and Conclusion

The results and observations show that both the price return and gross return for an LP follow a normal distribution centred at zero, which is true for more mature

Table 6.2: The table shows the VaR for price returns (V_p) and gross returns (V_g) using the Historical simulation method. Also shown is the difference between the two

Liquidity Pool description	**Value at Risk (VaR)**		
	price return (V_p)	**gross return** (V_g)	**difference** ($V_g - V_p$)
DAI-WETH Tick:10 Fee:0.5%	-0.01097	-0.01090	0.00007
HEX-USDC Tick:60 Fee:3%	-0.01146	-0.01135	0.00011
HEX-WETH Tick:60 Fee:3%	-0.01708	-0.01672	0.00036
SHIB-WETH Tick:60 Fee:3%	-0.01781	-0.01763	0.00018
USDC-WETH Tick:10 Fee:0.5%	-0.01039	-0.01017	0.00021
USDC-WETH Tick:60 Fee:3%	-0.01010	-0.00996	0.00014
WBTC-WETH Tick:10 Fee:0.5%	-0.01298	-0.01293	0.00005
WETH-LOOKS Tick:60 Fee:3%	-0.03039	-0.02948	0.00091
WETH-USDT Tick:10 Fee:0.5%	-0.01096	-0.01087	0.00009
WETH-USDT Tick:60 Fee:3%	-0.01002	-0.00996	0.00007

financial assets like stocks and cryptocurrencies. This shows that DeFi market behaviour shows a resemblance to any open financial market.

Furthermore, a key question, we need to address for this brief study was - should an LP provide liquidity to a DEX pool, or just hold the underlying ERC-20 tokens? The answer to this from a risk analysis perspective, based on the ten pools studied favours the former option, i.e. providing liquidity is less risky. We see this based on the estimation of VaR and ES using the historical simulation method. The gross returns have favourable VaR and ES as compared to the price returns, which is demonstrated by the positive difference of $V_g - V_p$ and $E_g - E_p$ (see 6.2 and 6.3). Furthermore, using the parametric method for risk estimation, we see a similar trend, except in a few pools where price returns seem to favour the gross returns as characterized by negative values of $V_g - V_p$ and $E_g - E_p$ (see tables 6.4 and 6.5 for details). Further investigation of these results on other DEX pools and platforms in currently ongoing.

Table 6.3: The table shows ES for price returns (E_p) and gross returns (E_g) using Historical Simulation method. Also shown is the difference between them

Liquidity Pool description	**Expected Shortfall (ES)**		
	price return (E_p)	**gross return** (E_g)	**difference** ($Eg - Ep$)
DAI-WETH Tick:10 Fee:0.5%	-0.019364	-0.019267	0.000097
HEX-USDC Tick:60 Fee:3%	-0.020899	-0.020576	0.000324
HEX-WETH Tick:60 Fee:3%	-0.027112	-0.026641	0.000471
SHIB-WETH Tick:60 Fee:3%	-0.028776	-0.028485	0.000291
USDC-WETH Tick:10 Fee:0.5%	-0.018221	-0.018012	0.000209
USDC-WETH Tick:60 Fee:3%	-0.017981	-0.017810	0.000170
WBTC-WETH Tick:10 Fee:0.5%	-0.021492	-0.021426	0.000066
WETH-LOOKS Tick:60 Fee:3%	-0.047798	-0.046653	0.001145
WETH-USDT Tick:10 Fee:0.5%	-0.019319	-0.019154	0.000165
WETH-USDT Tick:60 Fee:3%	-0.017553	-0.017392	0.000161

Table 6.4: The table shows the VaR for price returns (V_p) and gross returns (V_g) using the parametric variance-covariance method. Also shown is the difference between the two

Liquidity Pool description	**Value at Risk (VaR)**		
	price return (V_p)	**gross return** (V_g)	**difference** ($V_g - V_p$)
DAI-WETH Tick:10 Fee:0.5%	-0.009352	-0.009329	0.000023
HEX-USDC Tick:60 Fee:3%	-0.013731	-0.013503	0.000228
HEX-WETH Tick:60 Fee:3%	-0.014295	-0.013467	0.000828
SHIB-WETH Tick:60 Fee:3%	-0.016930	-0.017212	-0.000281
USDC-WETH Tick:10 Fee:0.5%	-0.008787	-0.008749	0.000037
USDC-WETH Tick:60 Fee:3%	-0.008866	-0.008856	0.000010
WBTC-WETH Tick:10 Fee:0.5%	-0.012346	-0.012328	0.000017
WETH-LOOKS Tick:60 Fee:3%	-0.029219	-0.029021	0.000198
WETH-USDT Tick:10 Fee:0.5%	-0.009835	-0.009784	0.000051
WETH-USDT Tick:60 Fee:3%	-0.009025	-0.008402	0.000624

Table 6.5: The table shows the ES for price returns (E_p) and gross returns (E_g) using the parametric variance-covariance method. Also shown is the difference between the two

Liquidity Pool description	**Expected Shortfall (ES)**		
	price return (E_p)	**gross return** (E_g)	**difference** ($E_g - E_p$)
DAI-WETH Tick:10 Fee:0.5%	-0.017240	-0.017285	-0.000045
HEX-USDC Tick:60 Fee:3%	-0.023766	-0.023358	0.000408
HEX-WETH Tick:60 Fee:3%	-0.024398	-0.023329	0.001070
SHIB-WETH Tick:60 Fee:3%	-0.027732	-0.028011	-0.000279
USDC-WETH Tick:10 Fee:0.5%	-0.016105	-0.016033	0.000072
USDC-WETH Tick:60 Fee:3%	-0.016353	-0.016264	0.000089
WBTC-WETH Tick:10 Fee:0.5%	-0.020681	-0.020644	0.000037
WETH-LOOKS Tick:60 Fee:3%	-0.046151	-0.045794	0.000358
WETH-USDT Tick:10 Fee:0.5%	-0.018041	-0.017906	0.000135
WETH-USDT Tick:60 Fee:3%	-0.016322	-0.015540	0.000783

Chapter 7

Conclusion and Contribution

This work studies the computational methods for market risk forecasting of Cryptocurrencies. These methods - primarily novel data-driven and neuro-volatility models are evaluated on cryptocurrencies and traditional stocks using fuzzy forecast intervals. We further applied this framework to forecast the risk for returns generated by algorithmic trading strategies.

We found that the proposed computational methods are effective in quantifying the risk, quantified in terms of key metrics like Value at Risk, Expected Shortfall, and Sharpe ratio by demonstrating narrower fuzzy forecasting intervals for risk. Furthermore, the methods found cryptocurrencies to have a significantly higher market risk as compared to stocks. However, when studying the algorithmic returns using a simple moving average crossover strategy, there was not any significant difference in risk between the two asset classes.

We also proposed a novel algorithmic trading strategy for high-frequency trading of cryptocurrencies, which consistently generates positive trade cashflows, along with having narrower fuzzy intervals as compared to a passive buy-and-hold strategy.

We applied our risk management models as a case study to DeFi ecosystem. We devoted this study to exploring the nascent DeFi ecosystem and applying risk management techniques to the returns generated by a Liquidity Provider in Decentralized Exchange. We conclude that the risk for a Liquidity Provider is reduced by providing liquidity to a DEX pool as compared to just holding the underlying tokens.

This thesis work can be further explored in a few directions, four of which are mentioned here:

- **Incorporating additional trading costs:** All the studies in this thesis assume zero trading fees and bid-ask spreads, which are unrealistic in real-world markets. We need to refine the models by incorporating these additional costs into asset returns. This may require an understanding of the trading costs charged by various asset exchanges along with advanced simulation models to factor in the bid-ask spread.

- **Improving the studies on minute-frequency data:** Due to the limited availability of high-frequency data on financial data platforms, our extended risk forecasting studies on the minute-frequency data could not reach a definitive conclusion. However, if data gets available for a longer time range, these studies can be expanded to reach some more meaningful conclusions regarding intra-day market risk and algorithmic trading.

- **DeFi ecosystem exploration:** Our study on the market risk in DeFi ecosystem is limited to a specific area of liquidity provision in decentralized exchange, while there are many other areas in DeFi investment which could be explored for market risk. Moreover, even the liquidity provision needs to be studied more extensively by considering a larger dataset from a wider set of liquidity pools considering more DEX platforms and multiple blockchains.

- **Fine-tuning the models for practicality in the industry:** A more rigorous backtesting of the models presented in this thesis can be done to deploy them for practical purposes in the investment management industry.

7.1 List of Contributions

We wrote the results from these investigations as multiple papers and published in most relevant venues as listed below:

7.1.1 Conference Publications

- **S. Bowala, J. Singh, A. Thavaneswaran, R. Thulasiram and S. Mandal, "Comparison of Fuzzy Risk Forecast Intervals for Cryptocurrencies, Proceedings of 2022 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr), 2022, pp. 1-8, doi: 10.1109/CIFEr52523.2022.9776213."**

  This paper uses fuzzy set theory along with data-driven volatility and neuro-volatility models to study fuzzy risk forecasts for cryptocurrencies and compare them to traditional technology stocks. The study focuses on long-term volatility forecasts with daily price data. (Chapter 5.1)

- **S. Dos Santos, J. Singh, R. K. Thulasiram, S. Kamali, L. Sirico and**

L. Loud, "A New Era of Blockchain-Powered Decentralized Finance (DeFi) - A Review", Proceedings of 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), 2022, pp. 1286-1292, doi: 10.1109/COMPSAC54236.2022.00203.

This paper reviews the key financial services offered in DeFi and draws a parallel to the corresponding services in the centralized financial industry. Some technical and economic risks associated with the DeFi investments are also discussed in the paper. This work aims to give an overview of the current state of the DeFi ecosystem in a lucid manner to make it accessible to a broader audience without compromising academic rigor. (Chapter 6)

- **J. Singh, S. Bowala, A. Thavaneswaran, R. Thulasiram and S. Mandal, "Data-Driven and Neuro-Volatility Fuzzy Forecasts for Cryptocurrencies", Proceedings of 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2022, pp. 1-8, doi: 10.1109/FUZZ-IEEE55066.2022.9882812.**

  In this paper, we study and compare the algorithmic trading returns of cryptocurrencies and stocks from a risk forecasting perspective using fuzzy forecasts of Sharpe ratio. A simple algorithmic trading approach, Simple Moving Average (SMA) crossover strategy, is used to calculate the algorithmic returns. (Chapter 5.3)

- **J. Singh, R. Thulasiram, A. Thavaneswaran, "LSTM based Algorithmic Trading model for Bitcoin", Proceedings of 2022 IEEE Symposium Series On Computational Intelligence (Forthcoming), Dec.**

2022, Singapore

This work explored the use of a Long Short Term Memory (LSTM) price prediction model to propose a novel algorithmic trading strategy for cryptocurrencies. The proposed novel high-frequency algorithmic trading strategy is tested for Bitcoin and Ethereum. This simple, yet effective trading algorithm uses the network's price forecasts to make buying and short-selling decisions for cryptocurrency based on certain set criteria. The proposed trading strategy gives positive returns when backtested on Bitcoin hourly prices taken from yahoo! finance. (Chapter 5.4)

7.1.2 Journal Publications

- **S.Bowala, J. Singh, "Optimizing Portfolio Risk of Cryptocurrencies Using Data-Driven Risk Measures", J. Risk Financial Manag. 2022, 15, 427. https://doi.org/10.3390/jrfm15100427**

  In this article, a data-driven portfolio risk measure is minimized to obtain the optimal portfolio weights. A recently proposed data-driven volatility forecasting approach with daily data is used to study risk forecasting for cryptocurrencies with high-frequency (hourly) big data. The paper emphasizes the superiority of the portfolio selection of cryptocurrencies by minimizing the recently proposed risk measure over the traditional minimum variance portfolio.

- **J. Singh, R. Thulasiram, A. Thavaneswaran, "Algorithmic and High-Frequency Trading in Cryptocurrency market - a Long Short-Term Model" (in preparation for submission to a journal)** This work is an extension of our earlier work which is to be presented in the 2022 IEEE Symposium Series On Computational Intelligence (Forthcoming), in Dec. 2022. We extend the analysis of our proposed algorithmic trading strategy to see how the underlying time-series forecasting models can impact the performance of the proposed strategy. We will emphasize the comparison of traditional models like ARIMA with deep-learning models as evaluated by the performance of strategy on the test dataset.

- **J. Singh, R. Thulasiram, A. Thavaneswaran, "Data-driven risk forecasting of intra-day cryptocurrency trading" (in preparation for submission to a journal)** This work is an extension of our earlier work presented at the 2022 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr). We are expanding the scope of study on a high-frequency intraday price dataset. The risk profile of intraday high-frequency trading is different in nature from long-term trading. We want to explore if our data-driven risk forecasting models work at this scale of trading data.

Bibliography

[1] "Basel III monitoring report," 2021. [Online]. Available: https://www.bis.org/bcbs/publ/d524.pdf

[2] "Understanding lstm networks." [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[3] P. F. Christoffersen and F. X. Diebold, "How relevant is volatility forecasting for financial risk management?" *Review of Economics and Statistics*, vol. 82, no. 1, pp. 12–22, 2000.

[4] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Business Review*, p. 21260, 2008.

[5] D. Fantazzini, E. Nigmatullin, V. Sukhanovskaya, and S. Ivliev, "Everything you always wanted to know about bitcoin modelling but were afraid to ask," *International Finance eJournal*, 2016.

[6] J. Danielsson, K. R. James, M. Valenzuela, and I. Zer, "Model risk of risk models," *Journal of Financial Stability*, vol. 23, pp. 79–91, 2016.

[7] A. Thavaneswaran, A. Paseka, and J. Frank, "Generalized value at risk forecasting," *Communications in Statistics - Theory and Methods*, vol. 49, no. 20, pp. 4988–4995, 2020.

[8] Y. Liang, A. Thavaneswaran, Z. Zhu, R. K. Thulasiram, and M. E. Hoque, "Data-driven adaptive regularized risk forecasting," in *2020 IEEE 44th Annual Computers, Software, and Applications (virtual) Conference (COMPSAC)*, 2020, pp. 1296–1301.

[9] A. Thavaneswaran, S. S. Appadoo, and A. Paseka, "Weighted possibilistic moments of fuzzy numbers with applications to garch modeling and option pricing," *Mathematical and Computer Modelling*, vol. 49, no. 1-2, pp. 352–368, 2009.

[10] A. Thavaneswaran, Y. Liang, Z. Zhu, and R. K. Thulasiram, "Novel data-driven fuzzy algorithmic volatility forecasting models with applications to algorithmic trading," in *2020 IEEE International (virtual) Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2020, pp. 1–8.

[11] S. Bodjanova, "Median value and median interval of a fuzzy number," *Information sciences*, vol. 172, no. 1-2, pp. 73–89, 2005.

[12] U. Cherubini, "Fuzzy measures and asset prices: accounting for information ambiguity," *Applied Mathematical Finance*, vol. 4, no. 3, pp. 135–149, 1997.

[13] A. L. Medaglia, S.-C. Fang, H. L. Nuttle, and J. R. Wilson, "An efficient and

flexible mechanism for constructing membership functions," *European Journal of Operational Research*, vol. 139, no. 1, pp. 84–95, 2002.

[14] S. Medasani, J. Kim, and R. Krishnapuram, "An overview of membership function generation techniques for pattern recognition," *International Journal of approximate reasoning*, vol. 19, no. 3-4, pp. 391–417, 1998.

[15] G. Meeden, "Fuzzy confidence intervals." [Online]. Available: http://users.stat.umn.edu/~gmeeden/talks/fuzz1.pdf

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[17] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[18] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[19] M. E. Hoque, A. Thavaneswaran, A. Paseka, and R. K. Thulasiram, "An algorithmic multiple trading strategy using data-driven random weights innovation volatility," in *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2021, pp. 1760–1765.

[20] W. Canny, "Deutsche bank survey shows most would hodl even

if crypto markets crashed," *CoinDesk Latest Headlines RSS*, Feb 2022. [Online]. Available: https://www.coindesk.com/business/2022/02/15/deutsche-bank-survey-shows-most-would-hodl-even-if-crypto-markets-crashed/

[21] S.-H. Poon and C. W. Granger, "Forecasting volatility in financial markets: A review," *Journal of Economic Literature*, vol. 41, no. 2, pp. 478–539, June 2003. [Online]. Available: https://www.aeaweb.org/articles?id=10.1257/002205103765762743

[22] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982. [Online]. Available: http://www.jstor.org/stable/1912773

[23] A. K. Bera and M. L. Higgins, "Arch models: Properties, estimation and testing," *Journal of Economic Surveys*, vol. 7, no. 4, pp. 305–366, 1993. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6419.1993.tb00170.x

[24] P. Jorion, "Risk2: Measuring the risk in value at risk," *Financial Analysts Journal*, vol. 52, no. 6, pp. 47–56, 1996. [Online]. Available: https://doi.org/10.2469/faj.v52.n6.2039

[25] A. Thavaneswaran, R. K. Thulasiram, Z. Zhu, M. E. Hoque, and N. Ravishanker, "Fuzzy value-at-risk forecasts using a novel data-driven neuro volatility predictive model," in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC), Milwaukee, WI, USA*, vol. 2. IEEE, 2019, pp. 221–226.

[26] A. Thavaneswaran, R. K. Thulasiram, J. Frank, Z. Zhu, and M. Singh, "Fuzzy option pricing using a novel data-driven feed forward neural network volatility model," in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), New Orleans, LA, USA*. IEEE, 2019, pp. 1–6.

[27] Z. Zhu, A. Thavaneswaran, A. Paseka, J. Frank, and R. Thulasiram, "Portfolio optimization using a novel data-driven ewma covariance model with big data," in *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 2020, pp. 1308–1313.

[28] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 2084–2123, 2016.

[29] G. M. Caporale and T. Zekokh, "Modelling volatility of cryptocurrencies using markov-switching garch models," *Research in International Business and Finance*, vol. 48, 12 2018.

[30] X. Sun, M. Liu, and Z. Sima, "A novel cryptocurrency price trend forecasting model based on lightgbm," *Finance Research Letters*, vol. 32, p. 101084, 2020.

[31] H. Sebastião and P. Godinho, "Forecasting and trading cryptocurrencies with machine learning under changing market conditions," *Financial Innovation*, vol. 7, no. 1, pp. 1–30, 2021.

[32] Y. Liu and A. Tsyvinski, "Risks and returns of cryptocurrency," *The Review of Financial Studies*, vol. 34, no. 6, pp. 2689–2727, 2021.

[33] S. A. Monfared and D. Enke, “Volatility forecasting using a hybrid gjr-garch neural network model,” *Procedia Computer Science*, vol. 36, pp. 246–253, 2014.

[34] W. Kristjanpoller and M. C. Minutolo, “Gold price volatility: A forecasting approach using the artificial neural network–garch model,” *Expert systems with applications*, vol. 42, no. 20, pp. 7245–7251, 2015.

[35] E. Rahimikia and S.-H. Poon, “Machine learning for realised volatility forecasting,” *Available at SSRN*, vol. 3707796, 2020.

[36] R. Huang and T. Polak, “Lobster: Limit order book reconstruction system,” *Available at SSRN 1977207*, 2011.

[37] J. Patterson, *Deep Learning: A Practitioner's Approach.* O'Reilly Media, 2017.

[38] S. Mishra, T. Ahmed, V. Mishra, M. Kaur, T. Martinetz, A. K. Jain, and H. Alshazly, “Multivariate and online prediction of closing price using kernel adaptive filtering,” *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–14, 2021.

[39] Y. Ning, L. Wah, and L. Erdan, “Stock price prediction based on error correction model and granger causality text,” *Cluster Computing*, vol. 2021, pp. 4849–4858, 2019.

[40] A. Hossain, R. Karim, R. K. Thulasiram, N. Bruce, and Y. Wang, “Hybrid deep learning model for stock price prediction,” in *Proc. 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Intl. Symp. on Com-*

putational Intelligence in Financial Engineering and Economics. IEEE doi: 10.1109/SSCI.2018.8628641., 2018, pp. 1837–1844.

[41] F. Glaser, K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering, "Bitcoin-asset or currency? revealing users' hidden intentions," *Revealing Users' Hidden Intentions (April 15, 2014). ECIS*, 2014.

[42] C. Baek and M. Elbeck, "Bitcoins as an investment or speculative vehicle? a first look," *Applied Economics Letters*, vol. 22, no. 1, pp. 30–34, 2015.

[43] J. R. Laboure, "The future of payments - part iii. digital currencies: The ultimate hard power tool," *Deutsche Bank Research*, Jan 2020.

[44] P. Katsiampa, "Volatility estimation for bitcoin: A comparison of garch models," *Economics Letters*, vol. 158, pp. 3–6, 2017.

[45] S. Corbet, A. Meegan, C. Larkin, B. Lucey, and L. Yarovaya, "Exploring the dynamic relationships between cryptocurrencies and other financial assets," *Economics Letters*, vol. 165, pp. 28–34, 2018.

[46] D. Vidal-Tomás and A. Ibañez, "Semi-strong efficiency of bitcoin," *Finance Research Letters*, vol. 27, pp. 259–265, 2018.

[47] A. Urquhart, "Price clustering in bitcoin," *Economics letters*, vol. 159, pp. 145–148, 2017.

[48] J. Rebane, I. Karlsson, P. Papapetrou, and S. Denic, "Seq2seq rnns and arima models for cryptocurrency prediction: A comparative study," in *SIGKDD Fintech'18, London, UK, August 19-23, 2018*, 2018.

[49] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, “A new forecasting framework for bitcoin price with lstm,” in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 2018, pp. 168–175.

[50] S. Alonso-Monsalve, A. L. Suárez-Cetrulo, A. Cervantes, and D. Quintana, “Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators,” *Expert Systems with Applications*, vol. 149, p. 113250, 2020.

[51] A. Arratia and A. X. López-Barrantes, “Do google trends forecast bitcoins? stylized facts and statistical evidence,” *Journal of Banking and Financial Technology*, vol. 5, no. 1, pp. 45–57, 2021.

[52] D. Shen, A. Urquhart, and P. Wang, “Does twitter predict bitcoin?” *Economics Letters*, vol. 174, pp. 118–122, 2019.

[53] R. J. Hyndman, G. Athanasopoulos, C. Bergmeir, G. Caceres, L. Chhay, M. O’Hara-Wild, F. Petropoulos, S. Razbash, and E. Wang, “Package ‘forecast’,” *Online] https://cran. r-project. org/web/packages/forecast/forecast. pdf*, 2020.

[54] D. J. Dubois, *Fuzzy sets and systems: theory and applications.* Academic press, 1980, vol. 144.

[55] H.-J. Zimmermann, *Fuzzy set theory—and its applications.* Springer Science & Business Media, 2011.

[56] A. Vantage, “Stock api, reimagined.” [Online]. Available: https://www.alphavantage.co/

[57] “Cryptocurrency prices, portfolio, forum, rankings.” [Online]. Available: https://www.cryptocompare.com/

[58] “Ethereum whitepaper,” (Accessed: 2022-09-24). [Online]. Available: https://ethereum.org/en/whitepaper/

[59] V. Buterin. Ethereum whitepaper — ethereum.org. Ethereum.org. [Online]. Available: https://ethereum.org/en/whitepaper/, last accessed 8 December 2021.

[60] S. Team. Solidity programming language. https://soliditylang.org/. [Online]. Available: https://soliditylang.org/, last accessed 8 December 2021.

[61] Binance smart chain - bsc. Binance. [Online]. Available https://dex-bin.bnbstatic.com/static/Whitepaper_%20Binance%20Smart%20Chain.pdf, last accessed 8 December 2021.

[62] Cardano project. Cardano.org. [Online]. Available https://cardano.org/, last accessed 8 December 2021.

[63] Solana project. Solana. [Online]. Available https://solana.com/, last accessed 8 December 2021.

[64] Avalanche project. Avalanche. [Online]. Available https://www.avax.network/, last accessed 8 December 2021.

[65] Defi pulse - the decentralized finance leaderboard. DefiPulse. [Online]. Available https://defipulse.com/, last accessed 8 December 2021.

[66] "Fluidefi - defi investments for institutional investors." [Online]. Available: https://fluidefi.com/

[67] Governance and funding — bank of england. [Online]. Available https://www.bankofengland.co.uk/about/governance-and-funding, last accessed 8 December 2021.

[68] Federal reserve board - the fed explained. [Online]. Available https://www.federalreserve.gov/aboutthefed/the-fed-explained.htm, last accessed 8 December 2021.

[69] Financial conduct authority. FCA. [Online]. Available https://www.fca.org.uk/about, last accessed 8 December 2021.

[70] Sec.gov — what we do. U.S Securities and Exchange Commission. [Online]. Available https://www.sec.gov/about/what-we-do, last accessed 8 December 2021.

[71] "Lecture 4 slides." [Online]. Available: https://rdi.berkeley.edu/berkeley-defi/assets/material/Lecture%204%20Slides.pdf

[72] U. W. Chohan, "Cryptocurrencies: A brief thematic review," *Available at SSRN 3024330*, 2017.

[73] S. Bano, A. Sonnino, M. Al-Bassam, S. Azouvi, P. McCorry, S. Meiklejohn, and G. Danezis, "Sok: Consensus in the age of blockchains," in *Proceedings*

of the 1st ACM Conference on Advances in Financial Technologies, 2019, pp. 183–198.

[74] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "Sok: Research perspectives and challenges for bitcoin and cryptocurrencies," in *2015 IEEE symposium on security and privacy*. IEEE, 2015, pp. 104–121.

[75] N. Szabo, "Formalizing and securing relationships on public networks," *First monday*, 1997.

[76] "Gas and fees ethereum.org," [Online]. Available https://ethereum.org/en/developers/docs/gas/#top, last accessed 8 December 2021.

[77] "Aave v2 whitepaper," Dec 2020. [Online]. Available: https://github.com/aave/protocol-v2/blob/master/aave-v2-whitepaper.pdf

[78] R. Leshner and G. Hayes, "Compound:the money market protocol," Feb 2019. [Online]. Available: https://compound.finance/documents/Compound.Whitepaper.pdf

[79] H. Adams, N. Zinsmeister, M. Salem, R. Keefer, and D. Robinson, "Uniswap v3 core," 2021.

[80] "Sushiswap, 2020. sushiswap staking," Dec 2020. [Online]. Available: https://docs.sushi.com/

[81] L. Kugler, "Non-fungible tokens and the future of art," *Communications of the ACM*, vol. 64, no. 9, pp. 19–20, 2021.

[82] S. Mishkin Frederic, "The economics of money, banking and financial markets," *Mishkin Frederic–Addison Wesley Longman*, 2004.

[83] L. Gudgeon, S. Werner, D. Perez, and W. J. Knottenbelt, "Defi protocols for loanable funds: Interest rates, liquidity and market efficiency," in *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, ser. AFT '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 92–112.

[84] M. Foundation, "The maker protocol: Makerdao's multi-collateral dai (mcd) system." [Online]. Available: https://makerdao.com/da/whitepaper/

[85] A. Finance, "Alpha homora v2." [Online]. Available: https://alphafinancelab.gitbook.io/alpha-finance-lab/alpha-products/3.-alpha-homora-v2-on-ethereum

[86] D. Wang, S. Wu, Z. Lin, L. Wu, X. Yuan, Y. Zhou, H. Wang, and K. Ren, "Towards a first step to understand flash loan and its applications in defi ecosystem," in *Proceedings of the Ninth International Workshop on Security in Blockchain and Cloud Computing*, 2021, pp. 23–28.

[87] D. e. a. Wang, "Towards a first step to understand flash loan and its applications in defi ecosystem," in *Proceedings of the Ninth International Workshop on Security in Blockchain and Cloud Computing*, ser. SBC '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 23–28.

[88] N. G. Mankiw, *Essentials of economics.* Cengage learning, 2020.

[89] P. Tasca, “The dual nature of bitcoin as payment network and money,” in *VI Chapter SUERF Conference Proceedings*, vol. 1, 2016.

[90] G. Hileman, “State of stablecoins (2019),” *Available at SSRN 3533143*, 2019.

[91] Klages-Mundt, “Stablecoins 2.0: Economic foundations and risk-based models,” in *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, 2020, pp. 59–79.

[92] D. Bullmann, “In search for stability in crypto-assets: are stablecoins the solution?” *ECB Occasional Paper*, no. 230, 2019.

[93] A. Gervais, “Decentralized exchanges (dex),” Sep 2021. [Online]. Available: https://berkeley-defi.github.io/assets/material/Updated%20Lecture%205%20Slides.pdf

[94] A. M. Othman, “Automated market making: Theory and practice,” Jun 2018. [Online]. Available: https://kilthub.cmu.edu/articles/thesis/Automated%5FMarket%5FMaking%5FTheory%5Fand%5FPractice/6714920/1

[95] G. Angeris and T. Chitra, “Improved price oracles,” *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, Oct 2020. [Online]. Available: http://dx.doi.org/10.1145/3419614.3423251

[96] K. Qin, L. Zhou, P. Gamito, P. Jovanovic, and A. Gervais, “An empirical study of defi liquidations: Incentives, risks, and instabilities,” *arXiv preprint arXiv:2106.06389*, 2021.

[97] T. Locke, "Over $10 billion was stolen in defi-related theft this year. here's how to protect yourself from common crypto scams," Dec 2021. [Online]. Available: https://www.cnbc.com/2021/12/14/common-defi-crypto-related-scams-and-how-to-protect-your-wallet.html

[98] S. Malwa, "Defi 'rug pull' scams pulled in $2.8b this year: Chainalysis," Dec 2021. [Online]. Available: https://www.coindesk.com/markets/2021/12/17/defi-rug-pull-scams-pulled-in-28b-this-year-chainalysis/

[99] "In the matter of zachary coburn, respondent. administrative proceeding file no. 3-18888," *Release No. 84553 / November 8, 2018*. [Online]. Available: https://www.sec.gov/litigation/admin/2018/34-84553.pdf

[100] J. Riley, "The current status of cryptocurrency regulation in china and its effect around the world," *China and WTO Review*, vol. 7, no. 1, pp. 135–152, 2021.

[101] "Crypto market cap charts," accessed on 09 Sept, 2022. [Online]. Available: https://www.coingecko.com/en/global-charts

[102] "Uniswap application," accessed on 09 Sept, 2022. [Online]. Available: https://app.uniswap.org/#/swap

[103] "Top decentralized exchanges on coingecko by trading volume," accessed on 09 Sept, 2022. [Online]. Available: https://www.coingecko.com/en/dex

[104] J. Hull and A. White, "Incorporating volatility updating into the historical simulation method for value-at-risk," *Journal of risk*, vol. 1, no. 1, pp. 5–19, 1998.

[105] D. Jadhav and T. Ramanathan, "Parametric and non-parametric estimation of value-at-risk," *The Journal of Risk Model Validation*, vol. 3, no. 1, p. 51, 2009.