Menemui Matematik (Discovering Mathematics) 47(1) (2025) 54-68

Menemui Matematik
(Discovering Mathematics)
journal homepage: https://persama.org.my/dismath/home

Risk Assessment In Cryptocurrencies: A Comparative Analysis Using Extreme Value Theory And Machine Learning for Predictive Insights

Rupika Selvaraja[1] and Wendy Ling Shinyie[2*]
[1,2]*Department of Mathematics and Statistics, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor*

[1]205892@student.upm.edu.my, [2]sy_ling@upm.edu.my
*Corresponding author

Received: 10 January 2025
Accepted: 10 April 2025

ABSTRACT
This study proposes a comprehensive approach to enhance risk management in financial market, especially in volatile cryptocurrency markets, by combining machine learning models and extreme event detection method. The main focus of the research is on analysing extreme occurrences in financial time series data, especially when it comes to the naturally volatile cryptocurrency markets. To detect extreme events, the research's initial phase uses the peaks over threshold approach to determine threshold values for six cryptocurrencies, namely, Bitcoin, Ethereum, Ripple, Flow, Solana, and Binance Coin. In order to predict future severe occurrences based on previous data, machine learning models such as support vector machine and random forest are used in the second stage of this study. The random forest model's ensemble approach is effective in capturing the complexities of cryptocurrency market dynamics and its versatility with different datasets and ability to identify intricate patterns make it a strong choice for forecasting severe occurrences in the future. Flow coin is found to be riskier than others with highest normalised value at risk and expected shortfall at all of the confidence levels. The integration of advanced statistical techniques, machine learning algorithms and risk management principles provides a robust framework for understanding and mitigating risks in cryptocurrency ventures. This research establishes a guide for further studies and advancements in the emerging field of digital asset investments.

Keywords: Extreme events, Cryptocurrencies, Machine learning

INTRODUCTION

Extreme value theory (EVT) offers a robust set of statistical tools designed to assess the probability and magnitude of extreme events that fall outside the typical data range. EVT encompasses various methods for analyzing and modeling rare events, including the block maxima method, peaks over threshold (POT), return periods, expected shortfall (ES), and value at risk (VaR). The POT method focuses on modelling exceedances over a specified high threshold, operating under the assumption that these exceedances are independent and identically distributed. VaR, widely used in the financial industry, is a key risk management tool that estimates potential financial losses during unexpected market conditions and quantifies the likelihood of such losses occurring. By integrating EVT into VaR calculations, financial institutions can better prepare for and mitigate the risks associated with extreme market movements.

Cryptocurrencies have garnered considerable interest and investment in financial markets due to their distinctive characteristics, which include decentralisation, transparency and limited availability. Nevertheless, the values of these electronic assets are subject to drastic changes and fluctuations, making them extremely unstable, which brings unusual challenges and dangers to investors and regulators. Hence, the conventional risk models and statistical models are insufficient to capture extreme events. Therefore, it is necessary to carry out research and utilises the concepts of EVT to scrutinise and formulate the model for the exceptional value patterns of cryptocurrencies. Besides that, when evaluating the risks of these digital assets, the potential losses and tail risks associated with cryptocurrency are important factors to be considered.

Cryptocurrency markets have long been characterised by extreme volatility, a feature that has drawn both attention and concern from investors and regulators alike. Historically, cryptocurrencies such as Bitcoin and Ethereum have experienced dramatic price swings within short periods. For instance, in 2017, Bitcoin surged to nearly $20,000 before crashing to below $6,000 within months, a pattern that repeated in 2021 when it again soared to over $60,000 before plummeting by more than 50% shortly after. These fluctuations highlight the susceptibility of cryptocurrencies to rapid market changes. Given the persistent volatility and the increasing adoption of cryptocurrencies, this research is crucial for developing robust risk management strategies. By leveraging EVT and machine learning models, this study aims to enhance the prediction of extreme market events and provide actionable insights to safeguard investors and stabilise markets in the face of these extraordinary risks.

Cryptocurrencies have attracted significant interest and investment in financial markets due to their unique features, including decentralisation, transparency, and limited supply. However, these digital assets are prone to extreme volatility, presenting substantial risks and challenges for both investors and regulators. Traditional risk and statistical models often fail to adequately capture these extreme fluctuations. Therefore, it is crucial to apply EVT in research to analyse and model the extreme value behavior of cryptocurrencies. Understanding potential losses and tail risks is essential when evaluating the overall risk profile of these digital assets. The primary objective of this research is to investigate tail risks and extreme events in cryptocurrency markets using EVT. The study will focus on applying machine learning models to forecast future extreme events in cryptocurrencies and employ EVT to calculate VaR and ES to assess potential losses and associated risks. Additionally, it will analyse the volatility and distributional properties of various cryptocurrencies, comparing them to provide insights. The research aims to offer actionable recommendations for investors, policymakers, and regulators to better manage and mitigate cryptocurrency risks. Machine learning (ML) complements traditional statistical methods like EVT in analysing complex financial systems, by handling non-linearities and offering strong predictive capabilities. Machine learning models, especially deep learning methods, can capture non-linear patterns and interactions that EVT might overlook, enhancing the analysis of tail risks in cryptocurrency markets. Moreover, they can better forecast future prices or volatility by learning from large datasets, complementing EVT's focus on extreme outcomes by providing more general market predictions. This study aims to fill the gap by integrating EVT with ML techniques to enhance the predictive power for extreme price movements and allowing more precise risk management strategies, which is crucial in cryptocurrency markets where existing literature often lacks robust risk assessment tools.

LITERATURE REVIEW

The peaks over threshold method was introduced when statisticians and hydrologists started researching extreme events, like floods and heavy rainfall, to determine their frequency and severity (de Fondeville and Davison, 2022). Gumbel (1958) first proposed the idea of fitting the POT data using generalised extreme value (GEV) distribution in 1958. This idea serves as the foundation for the POT method's analysis of extreme events. With the introduction of alternate distribution models like the generalised Pareto (GP) distribution and for the advancement of more complex statistical methods for analysing extreme events, the POT method has been further improved and expanded over time.

There have been several research papers that have looked into cryptocurrencies using the POT Method. Rui et al. (2022) present a dynamic POT Method to measure and forecast both the lower and upper tail VaR of Bitcoin returns. Their study illustrates how well the model predicts Bitcoin's tail risk and emphasises how it may be used in risk management and investment plans. Osterrieder and Lorenz (2017) studied the POT method's potential for examining the cryptocurrency market's tail behavior and its possible effects on risk control and investing tactics. Apart from that, by giving stylised information on their return and volatility qualities, the study adds to the body of knowledge on cryptocurrencies. It is also discovered that cryptocurrencies have significant skewness and volatility and that their returns have large tails and high kurtosis (Ghosh et al. 2023). The study also emphasises the need for more research in this field and the potential of cryptocurrencies as alternative financial assets.

The first neural network mathematical model was introduced in 1943, which launched the field of machine learning. The first computer learning program was originally written by Arthur Samuel in 1952, and it was designed to play checkers at the championship level. Afterward, Frank Rosenblatt created the perceptron, the first neural network intended specifically for computers, in 1957. There are different types of machine learning techniques, such as classification analysis, regression, data clustering, feature selection and extraction, dimensionality reduction, association rule learning, reinforcement learning, and deep learning techniques and Mahesh (2020) has offered a thorough analysis of how machine learning algorithms may be used to forecast Bitcoin values.

Risk measures in cryptocurrency are a necessary tool to allow better investment decisions and trading tactics. Likitracharoen et al. (2018) used historical and Gaussian parametric VaR to approximate VaR of several cryptocurrencies, including Bitcoin. Osterrieder and Lorenz (2017) and Gkillas and Katsiampa (2018) used EVT to estimate VaR. Stavroyiannis (2018), who used the generalised autoregressive conditionally heteroscedastic (GARCH) model and Pearson type-IV distribution to compute the VaR, is another example of related work. He achieved a decent performance for VaR 1% but a bad performance for both VaR 2.5% and 5%. Pele and Mazurencu-Marinescu-Pele (2019) projected a new method based on entropically to predict Bitcoin's daily VaR using high-frequency data, and they did so by comparing the proposed method to more established ones like historical, normal, and Student's t-GARCH (1,1). Furthermore, GJR-GARCH proposed by Glosten et al. (1993) was used to estimate VaR by considering the time-varying volatility of cryptocurrencies to estimate VaR considering regimes one, two, and three.

DATA AND METHODS USED

Data

This study focuses on the utilisation of six prominent cryptocurrencies: Binance Coin, Flow, Solana, Bitcoin, Ethereum, and Ripple. To obtain the necessary data, the price information was downloaded from the *coinmarketcap* website covering a specific sample period, spanning from 1 July 2022 to 1 July 2023, resulting in 366 closing prices for each respective series. After that, data adjustment process was carried out.

Peaks over threshold method

Threshold selection

Several methods are available for determining thresholds. This study employs the bootstrap method, a resampling technique that allows us to estimate population statistics from a sample. The core principle of the bootstrap method is to generate numerous new samples, known as bootstrap samples, to achieve a reliable threshold.

Single bootstrap procedure

A single bootstrap procedure was developed by Caeiro and Gomes (2015) to select the optimal threshold level. By overcoming the constrictive assumptions through the estimation of the necessary parameters, this procedure enhances the one introduced in Hall (1990). Using an auxiliary statistic, the bootstrap procedure simulates the Hill estimator's asymptotic mean square error (AMSE) criterion. The sample from a model is given as $X_n = X_1, X_2, X_3, \dots, X_n$. The semiparametric estimator is given as

$$\gamma_n(k) = \phi_n(X_n), \qquad 1 \le k < n.$$

The bootstrap sample from the model, F_n is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[X_i \le x]},$$

with $X_i = X_1, \dots, X_{n_1}, \quad n_1 \le n.$

The bootstrap estimator corresponding to the bootstrap sample is as below

$$\gamma_{n_1}(k_1) = \phi_{k_1}\left(X_{n_1}\right), \qquad 1 \le k_1 < n_1.$$

Hall proposed the minimisation of the bootstrap estimate of MSE, $\gamma_{n_1}(k_1)$ as follows

$$\mathrm{MSE}(n_1, k_1) = E[\{\gamma^*_{n_1}(k_1) - \gamma_{n_1}(k_0)\}^2 | X_n]$$

where $k_0 =$ the initial value of k such that $k_0 \to \infty$ as $n \to \infty$.

The bootstrap procedure will be applied to estimate the optimal sample fraction is

$$D^{\beta}_{n_1}\left(1 + o(1)\right)$$

with $\beta = \frac{2}{3}$ for both the original estimator and the bootstrap statistics. The optimal performance of the bootstrap estimator at a level k_1 where $\text{MSE}(\gamma^*_{n_1}(k_1)|X_n)$ is minimal is achieved.

$$k_0^*(n; k_0, n_1) = \frac{[k_{00}^*(n_1)]^2}{k_{00}^*(n_2)}$$

The threshold level, u is obtained when

$$\gamma_{n;n_1}(k_0) = \gamma_n(k_0^*(n; k_0, n_1)).$$

The AMSE of the Hill estimator is

$$\text{AMSE}(\gamma_n(k)) = E(\gamma_n(k) - \gamma)^2 = \text{Var}\big(\gamma_n(k)\big) + \text{Bias}^2\big(\gamma_n(k)\big).$$

Generalised Pareto distribution (GP)

To simulate the distribution of excesses over the selected threshold we used generalised Pareto distribution. Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed observations which is the closing price of the cryptocurrency with a common distribution function F and let

$$M_n = \max(X_1, X_2, \ldots, X_n).$$

The distribution M_n can be derived as:

$$P(M_n \le z) = P(X_1 \le z, \ldots, X_n \le z) = P(X_1 \le z) \ldots P(X_n \le z) = \big(F(z)\big)^n.$$

Parameter estimation

The scale parameter, σ and shape parameter, ξ, of the GP can be determined using maximum likelihood. Suppose that the values $y_1, y_2, \ldots, y_k$ are k excesses of threshold u. For $\xi \neq 0$ the log-likelihood which is derived from equation as

$$l(\sigma, \xi) = -k \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{k} \log\left(1 + \frac{\xi}{\sigma} y_i\right)$$

where $(1+\frac{\xi}{\sigma}y_i\ )>0$ for $i=1,\ldots,k$; otherwise, $l(\sigma,\xi)=-\infty$. In the case, $\xi=0$ the log-likelihood is derived from as

$$l(\sigma)=-k\log\sigma\ -\frac{1}{\sigma}\sum_{i=1}^{k}y_i$$

Support vector machine (SVM)

Radial basis function (RBF) kernel

The RBF kernel is commonly referred to as the Gaussian kernel, for datasets containing intricate and non-linear patterns or clusters. A localised and finite reaction is possible throughout the whole x-axis by mapping the input data into an infinite-dimensional space. The RBF kernel is defined as follows

$$K(x,y)=\exp(-\gamma|x-y|^2),$$

γ as a positive parameter that controls the shape of the kernel.

Optimal hyperplane for SVM model

Lagrangian formulation is applied in nonlinear SVM models to determine the ideal hyperplane in higher-dimensional space that divides data points of various classes. With nonlinear SVM models, the Lagrangian formulation is very helpful as it enables the data to be transformed into a higher-dimensional space using kernel functions, perhaps leading to a linear separation. The Lagrange multipliers obtained during this process are then used in the decision function of the SVM, which is a key component of the model's ability to classify and make predictions based on the input data. The decision function of the SVM is as follows

$$f(x)=\sum_{i=1}^{n}a_iy_iK(x,y)+b$$

where

$f(x)$	: decision function
a_i	: the Lagrange multipliers obtained during the training process
y_i	: class label
$K(x,y)$	: RBF kernel
b	: bias term.

Model training and prediction

The SVM model must be trained before predictions can be made. The best hyperplane is found by the algorithm by adjusting parameters during training. The trained SVM is then applied to test data and subsequent data points. The accuracy and precision of the model are assessed by the computation of performance measures like root mean squared error. Then the trained SVM model is used to forecast future extreme events in the cryptocurrency markets, offering insightful information.

Random forest model

Data splitting

The dataset was divided into two parts which were the training set and testing set. This is usually done to assess the model's performance on an additional, unseen subset after training it on one. The dataset D in the context of cryptocurrencies comprises extreme events of the closing price of the cryptocurrencies. The data for each observation i was represented as (X_i, Y_i) where X_i (extreme events from the historical data) is a vector of predictor variables and Y_i (future extreme events) is the target variable. The data splitting process is represented mathematically as follows:

$$D = D_{train} \cup D_{test}$$

where D_{train} : training set which contains 80% of the data
D_{test} : test data contains 20% of the data.

Random forest model training

The random forest model, ensemble N decision trees (e.g: $N = 140$).

$$\text{RF}_{\text{model}} = \{T_1, T_2, \ldots, T_n\}$$

Each decision tree T_i is trained using a bootstrapped subset of the training data D_{train} with a random subset of features examined at each split.

$$T_i = \text{Train Decision Tree } (D_{\text{train},i})$$

The predictions of future extreme events are obtained by aggregating the predictions of the individual decision trees,

$$\text{predictions } = \frac{1}{N}\sum_{i=1}^{N} \text{Predict}(T_i, D_{\text{test}}).$$

The outputs of individual decision trees are averaged to provide predictions, which results in a more reliable and accurate model.

Root mean square error (RMSE) and relative root mean square error (RRMSE)

This analysis involves a simulation that compares the predictive performance of the support vector machine and random forest model using the root mean square error and relative root mean square error. These indicators show how well the models work by gauging the precision of predictions made in a regression setting. RMSE is derived as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

where
n : the number of data points
y_i : the actual extreme event at i^{th} point
$\hat{y}_i$: the predicted extreme event at i^{th} point, and RRMSE is

$$\text{RRMSE} = \left(\frac{\text{RMSE}}{\text{Mean}|y|}\right) \times 100$$

where Mean$|y|$ is the mean of the absolute values of the actual extreme events.

Risk Measure

Value at Risk (VaR)

VaR is a statistical method that is used to calculate the potential loss of a specific investment or a portfolio of investments over a certain period, at a specific level of confidence under normal market conditions. The formula that will be used to calculate VaR using GP is as follows:

$$\text{VaR} = u + \frac{\sigma}{\xi}\left[\left(\frac{n}{N_u}a\right)^{-\xi} - 1\right]$$

where N_u means the number of observations that exceeded the given threshold.

Expected shortfall (ES)

ES which is also known as the conditional value at risk quantifies the amount of tail risk present in an investment portfolio. The expected average loss that an investment can experience above a specific confidence level is denoted by the expected shortfall

$$\mathrm{ES} = \int_a^1 u + \frac{\sigma}{\xi}\left[\left(\frac{n}{N_u}a\right)^{-\xi} - 1\right]dx = \frac{\mathrm{VaR} + \sigma - u\xi}{1-\xi}, \text{for } \xi \neq 0.$$

RESULTS AND DISCUSSION

Threshold level

The threshold levels determined for each cryptocurrency using single bootstrap procedure are shown in Table 1. The threshold of Bitcoin which is 29340.26 shows its role in determining the extreme events in the closing prices of Bitcoin. As a result, sudden price spikes or sharp declines occur. The threshold values vary greatly in size among various cryptocurrencies. For example, compared to Flow and Ripple, the threshold values for Bitcoin and Ethereum are significantly greater.

Table 1: Threshold levels for cryptocurrencies

Cryptocurrencies	Thresholds
Bitcoin (BTC)	29340.3
Ethereum (ETH)	1884.5
Ripple (XRP)	0.5061
Flow (FLOW)	1.8616
Solana (SOL)	36.7658
Binance Coin (BNB)	324.91

Parameter estimation

The generalised Pareto distribution (GP) model was used to describe the characteristics of the extreme events in the dataset. Based on the observed extreme events, we determined a threshold for each cryptocurrency and fitted the GP to exceedances over this threshold using the maximum likelihood estimation (MLE) method. As a result, this allowed us to estimate the two parameters which are the shape and the scale parameters. This has provided insight into the tail behavior and spread or the width of the distribution. The results that explained the statistical properties of the cryptocurrency's price distributions will offer a nuanced perspective on risk management, volatility, and more. The estimation of the parameters is presented in Table 2.

Table 2: Parameter estimates for cryptocurrencies

Cryptocurrencies	Shape Parameter, ξ	Scale Parameter, σ
Bitcoin	-1.6274	2205.51
Ethereum	0.2908	44.327
Ripple	-0.7030	0.0269

Flow	−0.2720	0.5393
Solana	−0.7861	7.7081
Binance	−0.2330	10.189

For Bitcoin, a negative shape parameter indicates a heavy-tailed distribution, suggesting a higher likelihood of extreme events due to the slow decline in the tail. The high scale parameter reflects that these extreme events are widely distributed in magnitude. For Ethereum, the positive shape parameter implies that the tail is not significantly heavy, indicating less extreme behavior. The scale parameter shows variability in extreme events, with magnitudes spread over a range. Ripple's negative shape parameter points to a heavy right tail, while the scale parameter indicates a significant spread in the data. Flow's negative shape parameter deviates from the norm, showing a statistically significant heavy tail, and the high scale value reflects larger magnitudes. Solana also exhibits heavy-tailed behavior with more extreme events than expected in a standard distribution. The high scale parameter suggests a broader range of magnitudes. Finally, Binance Coin's negative shape parameter suggests a heavy right tail, and the large scale parameter indicates significant variability in the data. Cryptocurrencies with negative shape parameters like Bitcoin, Ripple, Flow, Solana, and Binance exhibit a heavy-tailed behavior suggesting a greater chance of extreme events. Furthermore, high-scale parameters (Bitcoin, Solana, Binance) imply that extreme events are dispersed over large-magnitude distributions. Risk management requires an understanding of magnitude variability and tail behavior. Overall, these parameter estimates help with risk assessment and management techniques by offering insights into the magnitude variability and tail behavior of extreme events for each cryptocurrency. Awareness of these traits aids in creating risk models and mitigation techniques that are more reliable and customised to the unique features of each coin.

Model performance

Support vector machine (SVM) and random forest models were used to predict future extreme events based on the peaks obtained through the POT method. Table 3 and 4 summarises the RMSE and RRMSE values for both models across six cryptocurrencies.

Table 3: RMSE and RRMSE results for support vector machine model

Cryptocurrency	Bitcoin	Ethereum	Ripple	Flow	Solana	Binance
RMSE	7739.98	247.43	0.1186	1.3954	33.108	84.731
RRMSE	0.5226	0.5043	0.3783	6.728	0.4007	0.9756

Table 4: RMSE and RRMSE results for random forest model

Cryptocurrency	Bitcoin	Ethereum	Ripple	Flow	Solana	Binance
RMSE	7789.52	246.73	0.1162	1.418	33.051	82.446
RRMSE	0.5259	0.5029	0.3710	6.837	0.4000	0.9730

For Bitcoin both the models have similar RMSE and RRMSE values. The RRMSE of both models is around 0.52. This indicates a moderate level of predictive error. Hence, the SVM and the random forest model both perform similarly for BitCoin. For Ethereum, the RMSE is similar for both the SVM and random forest model but the RRMSE for the random forest model is better than the SVM model. For Ripple, both models have very low RMSE value, which shows a strong predictive analysis. The RRMSE of the random forest model is also lesser compared to the SVM Model.

As for Flow, the RRMSE is large, indicating that both models have greater difficulty forecasting severe occurrences than for other cryptocurrencies. The random forest model shows slightly better RRMSE for Flow. Furthermore, for Solana, the SVM's RRMSE is slightly smaller than random forest model's. The SVM has slightly better relative RMSE for Binance compared to random forest model.

In conclusion, with low RRMSE and low RMSE values, both models function well for Ripple. Both models perform similarly for Ethereum and Bitcoin. With comparatively high RRMSE values for both models, the Flow appears to be the cryptocurrency that is most difficult to be predicted. Based on the better RRMSE values, the random forest model seems to have an advantage over Ethereum, Ripple, Solana and Binance. This indicates that random forest model is better at handling noisy and complex data by averaging out predictions from multiple decision trees and make it more robust for cryptocurrencies with more volatile or complex price movements.

Risk evaluation

Table 5: Value at risk results for six cryptocurrencies

Cryptocurrencies	1% VaR	5% VaR	10% VaR
Bitcoin	1.0439	1.0143	0.9478
Ethereum	1.0738	1.0160	0.9983
Ripple	1.0602	1.0277	0.9976
Flow	1.5236	1.2261	1.0520
Solana	1.2268	1.1252	1.0228
Binance	1.0570	1.0218	1.0020

Table 6: Expected shortfall results for six cryptocurrencies

Cryptocurrencies	1% ES	5% ES	10% ES
Bitcoin	1.045307	1.034071	1.008748
Ethereum	1.137177	1.055677	1.030748
Ripple	1.06658	1.04750	1.0298
Flow	1.639358	1.405513	1.268662

Solana	1.244359	1.187506	1.130131
Binance	1.07169	1.043075	1.027036

Table 5 and 6 presents the value at risk (VaR) and expected shortfall (ES) values calculated at three different confidence levels which are 1%, 5%, and 10% respectively. At all the confidence levels, Flow Coin has the highest normalised VaR and ES. This shows that it is riskier than the others. Bitcoin continuously displays lower normalised VaR and ES indicating a reduced level of risk. The risk associated with Solana is higher, specifically at a 1% confidence interval, suggesting a larger downside risk. At varying degrees of certainty, the risk profiles of Ethereum and Binance are comparable.

Several risk management techniques, based on the normalised VaR and ES analysis, are available to investors navigating the cryptocurrency market. The key to reducing the impact of individual asset volatility is diversification which involves spreading the investments over a variety of cryptocurrencies. Regular monitoring of market conditions, news, and regulatory development is very important for decision-making. Investors should assess their risk tolerance and align it with their chosen investment strategy, adopting a more conservative approach. Furthermore, it is critical to maintain up-to-date knowledge of the unique traits, applications, and possible risks associated with each of the cryptocurrencies.

Active volatility management, hedging strategies, and caution in portfolio exposure are recommended for assets with higher risk profiles, such as Solana or Flow. Conversely, stable assets like Bitcoin can serve as a stabilising element in a diversified portfolio, particularly for those with a lower risk appetite. In summary, managing the ever-changing world of cryptocurrency investments requires a well-rounded and knowledgeable approach, as well as constant observation and flexibility.

Volatility of the cryptocurrencies

Annualised volatility is the degree of price swings or variability in an asset's daily returns over a given time frame, usually a year. It provides information on the possible size of price fluctuations, which helps to quantify the amount of risk involved with an investment. The annualised volatility of the six cryptocurrencies is shown in Table 7.

Table 7: Annualised volatility for cryptocurrencies

Cryptocurrencies	Annualised Volatility
Bitcoin	0.5245
Ethereum	1.1289
Ripple	1.0207
Flow	0.7076
Solana	0.5691
Binance Coin	0.7140

Out of all the cryptocurrencies analysed in this study, Ethereum has the greatest annualised volatility, meaning that its price swings are the largest. When compared to the other cryptocurrencies, Bitcoin has the lowest annualised volatility, indicating more steady price changes. The modest volatility levels of Flow, Solana and Binance Coin show a balance between price fluctuations and stability, while Ripple exhibit moderate to high volatility.

As shown with Ethereum, more volatility might present chances for significant gains. On the other hand, it also increases risk because prices might change quickly. Those who are looking for these kinds of assets should have risk management plans in place and be well-prepared for the volatility that comes with them. Conversely, cryptocurrencies with lower annualised volatility like Bitcoin and Solana are thought to be more stable. Investors who value consistency and are wary of potentially significant price fluctuations could find this reduced volatility appealing. Such investments might act as the cornerstone of a more cautious strategy.

Investors need to understand that the marketplaces for cryptocurrencies are quite volatile. It is important to consistently observe market circumstances and make necessary modifications to investing strategy. The state of the market, changes in regulations, and breakthroughs in technology are just a few of the variables that can cause volatility levels to fluctuate over time. One essential tactic for reducing risk is still diversification. Effective portfolio risk management may be achieved by having a well-diversified cryptocurrency portfolio that consists of assets with various risk profiles. Volatility may present short-term trading opportunities, but maintaining a long-term outlook is essential for long-term success in the cryptocurrency space. Comprehending the underlying principles that propel the worth of a cryptocurrency and remaining updated on prospective advancements may enhance the scope of an investing plan.

CONCLUSION

In this study, we explored extreme events in the cryptocurrency markets, focusing on six prominent cryptocurrencies: Bitcoin, Ethereum, Ripple, Flow, Solana, and Binance. By using the peaks over threshold approach and machine learning models (support vector machine and random forest), we aimed to predict future extreme events. Our analysis showed that the random forest model consistently had lower RRMSE values than SVM, suggesting superior predictive capabilities. Notably, Ethereum, Ripple, Solana, and Binance displayed lower relative RRMSE values within the random forest model, indicating a better capacity to forecast severe occurrences. The model's ensemble approach captures the complexities of cryptocurrency market dynamics, making it a strong choice for future forecasting. We also examined risk management through volatility, value at risk (VaR), and expected shortfall (ES). Flow coin emerged as the riskiest, with the highest normalized VaR and ES across all confidence levels, emphasizing the importance of understanding and mitigating potential losses.

The practical implications of this findings on cryptocurrency trading and investment strategies are significant in order to provide insights into which model performs better under specific conditions, allowing traders to select the most suitable model for predicting price movements. This can lead to more accurate trading decisions and reduced financial risk. Besides that, traders can better time their entry and exit points, potentially increasing profits and minimizing losses. This is especially crucial in the highly volatile cryptocurrency market.

The limitation of this study is the challenge of generalising findings across different cryptocurrencies due to their unique market behaviors and varying degrees of volatility. This makes it difficult to create universally applicable models that effectively predict price movements or manage risk across all digital assets. Future research could explore adaptive machine learning models that can dynamically adjust to the specific characteristics of each cryptocurrency. Additionally, investigating the integration of real-time data streams and external factors such as regulatory changes and macroeconomic indicators into time series models may provide a more comprehensive understanding of cryptocurrency risk management.

In conclusion, the amalgamation of sophisticated statistical methodologies, machine learning strategies, and risk management tenets furnishes a sturdy structure for comprehending and alleviating hazards within the domain of cryptocurrency. This research provides a strong basis for future exploration and innovation in the emerging field of digital asset investments.

REFERENCES

Caeiro, J. and Gomes, M. I. (2015), Threshold selection in extreme value analysis. *Extreme Value Modeling and Risk Analysis:Methods and Applications*, 69 – 87.

de Fondeville, R. and Davison, A. C. (2022), Functional peaks-over-threshold analysis. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **84(4)**: 1392 – 1422.

Ghosh, B., Bouri, E., Wee, J. B. and Zulfiqar, N. (2023), Return and volatility properties: Stylized facts from the universe of cryptocurrencies and NFTs. *Research in International Business and Finance*, **65**, 101945.

Gkillas, K. and Katsiampa, P. (2018), An Application of Extreme Value Theory to Cryptocurrencies. *Economics Letters*, **164**: 109 – 111.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance,* **48(5):** 1779 – 1801.

Gumbel, E. J. (1958), *Statistics of Extremes*. New York: Columbia University Press.

Hall, P. (1990), Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, **32(2)**: 177 – 203.

Rui, K., Luyao, Y. and Changchun, T. (2022), Forecasting tail risk for Bitcoin: A dynamic peak over threshold approach. *Finance Research Letters*, **49**, 103086.

Likitratcharoen, D., Ranong, T. N., Chuengsuksomboon, R., Sritanee, N. and Pansriwong, A. (2018), Value at Risk Performance in Cryptocurrencies. *The Journal of Risk Management and Insurance*, **22(1)**: 11 – 28.

Mahesh, B. (2020), Machine Learning Algorithms - A Review. *International Journal of Science and Research*, **9(1)**: 381 – 386.

Osterrieder, J. and Lorenz, J. (2017), A Statistical Risk Assessment of Bitcoin and Its Extreme Tail Behavior. *Annals of Financial Economics*, **12(01)**: 1750003.

Pele, D. T. and Mazurencu-Marinescu-Pele, M. (2019), Using High-Frequency Entropy to Forecast Bitcoin's Daily Value at Risk. *Entropy*, **21(2)**: 102.

Stavroyiannis, S. (2018), Value-at-Risk and Related Measures for the Bitcoin. *The Journal of Risk Finance*, **19(2):** 127 – 136.