Scenario Generation and Stress Testing for Cryptocurrency Markets using GAN and Diffusion-Based Generative Models

Hassan Oukhouya[†‡*], **Zakaria Moussaoui**[§*], **Moustapha Faizi**[†], **Raby Guerbaz**[‡], **Badr Abou El Majd**[§]
[†]LaMSD, MSASE, FSJES, Mohammed First University of Oujda, Morocco
[‡]MAEGE, FSJES Ain Sebaâ, Hassan II University of Casablanca, Morocco
[§]LMSA Laboratory, Faculty of Science, Mohammed V University in Rabat, Morocco
[*]oukhouya.hassan@ump.ac.ma,zakaria.moussaoui@um5r.ac.ma

Abstract

This study presents a hybrid generative framework that combines diffusion models and GANs to generate realistic cryptocurrency market scenarios under both normal and stressed conditions. Applied to significant Layer-1 assets (BTC, ETH, XRP, LTC), the model incorporates macroeconomic and global financial indicators to capture structural and cyclical dynamics in highly volatile, data-scarce environments. Compared to traditional forecasting and simulation methods, the hybrid approach enhances stress testing robustness and interpretability, supported by SHAP-based analysis, providing valuable insights for financial risk management and informed portfolio decision-making.

Key words: Diffusion Models; Generative Adversarial Networks (GANs); Scenario Generation; Stress Testing; Cryptocurrency Markets.

1 Introduction

Machine learning (ML) has become increasingly important for analyzing financial time series, providing enhanced insights for risk management and informed decision-making. However, economic data remain difficult to model due to limited sample size, short historical coverage, and low signal-to-noise ratios, which constrain the effectiveness of deep learning (DL) methods [1]. In addition, financial series exhibit stylized facts-such as fat-tailed returns, volatility clustering, and seasonal patterns-that complicate the generation of realistic synthetic data [2]. Recent advances in generative modeling, including generative adversarial networks (GANs), variational autoencoders (VAEs), and diffusion models, have shown strong potential for capturing complex distributions across multiple data modalities [3]. In particular, diffusion models have emerged as state-of-the-art tools for time-series imputation and the generation of synthetic data. Despite their promise, generative models remain underutilized in financial time series due to data scarcity and the lack of domain-specific constraints, which often lead to unstable training outcomes. Yet, they offer substantial potential for enhancing scenario generation and stress testing [4], particularly in data-constrained environments such as those in the Middle East and North Africa (MENA) region, including financial markets [5, 6]. To address these limitations, this study investigates diffusion- and GAN-based frameworks for generating financial scenarios, aiming to improve model robustness, interpretability, and real-world relevance in stress-testing applications. Recent advances in generative modeling have significantly improved scenario generation and stress testing for volatile financial assets, including cryptocurrencies. GANs have demonstrated strong predictive performance for Bitcoin using on-chain and market indicators [7]. At the same time, diffusion-based approaches have shown the ability to reproduce key stylized

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

facts such as fat tails and volatility clustering [2]. Several diffusion frameworks—such as models integrating geometric Brownian motion [8], FTS-Diffusion for irregular multi-scale patterns [9], and controllable financial diffusion systems like CoFinDiff [10] — have demonstrated improved fidelity in synthetic financial time series. Hybrid architectures combining GANs and diffusion processes, including Diffusion-GAN [11, 12] and recent controllable hybrid models [13], further enhance data efficiency and stability. Although these developments highlight the growing potential of generative models in finance, few studies focus specifically on hybrid modeling for cryptocurrency stress testing. Our work addresses this gap by integrating GAN and diffusion methods with SHAP and LIME explainability, and evaluating their performance across major Layer-1 cryptocurrencies. The remainder of this paper is structured as follows. Section 2 outlines the methodology, including the dataset description, preprocessing steps, and the predictive models employed. Section 3 presents the experimental results and provides a detailed discussion of the comparative performance of individual and hybrid approaches. Finally, Section 4 concludes the study by summarizing the key findings and highlighting potential directions for future research.

2 Methodology

The empirical analysis uses historical time series data from four major cryptocurrencies-Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), and Litecoin (LTC)-chosen for their prominence and heterogeneous volatility structures, which make them strong benchmarks for evaluating generative financial models [14–16]. The training workflow used in this study is outlined in Algorithm 1, which details the CGAN optimization process employed for generating financial scenarios. Table 1 presents the

Algorithm 1: Training Procedure for Conditional GAN (CGAN)

Input: Training dataset $\mathcal{D} = \{x_i\}_{i=1}^{N}$, sequence length L, latent dimension $z \in \mathbb{R}^d$, learning rate η, batch size B, number of epochs E

Output: Trained Generator G_θ and Discriminator D_ϕ

Initialize Generator $G_\theta : \mathbb{R}^d \to \mathbb{R}^L$ and Discriminator $D_\phi : \mathbb{R}^L \to [0, 1]$

Initialize optimizers $\text{Adam}_G(\theta, \eta)$ and $\text{Adam}_D(\phi, \eta)$

for *epoch* $= 1$ **to** E **do**

  for *each minibatch* $\{x^{(1)}, \dots, x^{(B)}\}$ *from* $\mathcal{D}$ **do**

    Sample latent noise $z^{(i)} \sim \mathcal{N}(0, I_d)$ for $i = 1, \dots, B$

    Generate fake samples: $\tilde{x}^{(i)} = G_\theta(z^{(i)})$

    Compute discriminator loss:

$$\mathcal{L}_D = -\frac{1}{B} \sum_{i=1}^{B} \Big[\log D_\phi(x^{(i)}) + \log(1 - D_\phi(\tilde{x}^{(i)})) \Big]$$

    Update discriminator: $\phi \leftarrow \phi - \eta \nabla_\phi \mathcal{L}_D$

    Resample latent noise $z^{(i)} \sim \mathcal{N}(0, I_d)$

    Generate new fake samples: $\tilde{x}^{(i)} = G_\theta(z^{(i)})$

    Compute generator loss:

$$\mathcal{L}_G = -\frac{1}{B} \sum_{i=1}^{B} \log D_\phi(\tilde{x}^{(i)})$$

    Update generator: $\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_G$

  end

  Report average $\mathcal{L}_D, \mathcal{L}_G$ for monitoring

end

return G_θ, D_ϕ

main dataset specifications, including the study period, sequence length, total observations, and the train–test division employed for model development and assessment. Conditional Generative Adversarial Networks (CGANs) build upon the original GAN formulation of [17] by incorporating auxiliary conditioning variables into both the generator and discriminator, enabling more controlled and targeted data synthesis [18]. The pseudo-algorithm 2 outlines a simplified denoising diffusion framework implemented with a multilayer perceptron (MLP). In this approach, Gaussian noise is added to input sequences during training, and the model learns to reconstruct the clean signal by

Table 1: Description of the cryptocurrency time series dataset used in this study.

Asset	Ticker	Period	Seq. Length	Total Obs.	Train	Test
Bitcoin	BTC-USD	2015-08-01 – 2025-08-01	30	3622	2897	725
Ethereum	ETH-USD	2015-08-01 – 2025-08-01	30	3622	2897	725
Ripple	XRP-USD	2015-08-01 – 2025-08-01	30	3622	2897	725
Litecoin	LTC-USD	2015-08-01 – 2025-08-01	30	3622	2897	725

minimizing a mean-squared error loss. Each epoch involves corrupting mini-batches, denoising them through the network, and updating parameters using the Adam optimizer. This procedure follows the principles of diffusion probabilistic models [19] and denoising diffusion probabilistic models (DDPMs) [20], which have shown strong generative performance in high-dimensional settings. To

Algorithm 2: Simple Diffusion Denoiser (Prototype)

Input: Training data $\mathcal{D} = \{x_i \in \mathbb{R}^L\}_{i=1}^N$, sequence length L, noise scale σ, learning rate η, epochs E, batch size B.
Output: Trained denoising network $f_\theta : \mathbb{R}^L \to \mathbb{R}^L$
Initialize network parameters θ of $f_\theta(x) = W_2\,\sigma(W_1 x + b_1) + b_2$ (MLP).
Initialize optimizer Adam with learning rate η.
for $epoch \leftarrow 1$ **to** E **do**
  Shuffle $\mathcal{D}$ into mini-batches $\{x^{(b)}\}$ of size B.
  for *each batch* $x^{(b)}$ **do**
    Add Gaussian noise: $\tilde{x}^{(b)} = x^{(b)} + \epsilon^{(b)}$, with $\epsilon^{(b)} \sim \mathcal{N}(0, \sigma^2 I)$.
    Forward pass: $\hat{x}^{(b)} \leftarrow f_\theta(\tilde{x}^{(b)})$.
    Compute loss: $\mathcal{L}^{(b)}(\theta) = \frac{1}{B}\sum_{i=1}^{B} \|\hat{x}_i^{(b)} - x_i^{(b)}\|^2$.
    Backpropagate gradients $\nabla_\theta \mathcal{L}^{(b)}(\theta)$.
    Update $\theta \leftarrow \theta - \eta \cdot Adam(\nabla_\theta \mathcal{L}^{(b)}(\theta))$.
  Record epoch loss $\mathcal{L}_{epoch}$.
return trained model f_θ and training history.

enhance the realism and variability of generated financial time series, we integrate GANs with diffusion models within a unified hybrid framework. GANs effectively learn complex data distributions through adversarial training [17], whereas diffusion models excel at producing high-fidelity samples by progressively denoising noise [20, 21]. Integrating these two approaches enables the hybrid method to benefit from the expressive generative capacity of GANs and the stability and resistance to mode collapse of diffusion processes. This ensemble yields more stable, diverse, and realistic synthetic sequences suitable for financial simulation and data augmentation. The whole procedure is detailed in Algorithm 3. The workflow begins with preprocessing cryptocurrency time series into return-based sequences, which better represent market dynamics. Diffusion and GAN models are trained independently to capture complementary statistical properties of the data, and their outputs are then combined into a hybrid generator that integrates the stability of diffusion mechanisms with the flexibility of GAN learning. The resulting synthetic scenarios undergo interpretability analysis using SHAP and LIME to assess the economic plausibility of generated patterns. Model performance is then evaluated through RMSE, MAE, and MAPE, ensuring that the framework meets predictive and financial relevance criteria.

3 Results & Discussion

The empirical results indicate that GAN-based models capture short-term dependencies but struggle with longer-range temporal structure, a limitation consistent with prior work on adversarial learning. Diffusion models improve the reproduction of autocorrelation patterns, particularly for assets with smoother volatility dynamics such as ETH-USD and XRP-USD. The hybrid GAN–Diffusion framework provides the most reliable performance, yielding more stable and realistic dependence structures across all cryptocurrencies. This aligns with previous evidence showing that combining adversarial and diffusion mechanisms enhances generative robustness [22, 23, 13]. Quantitatively,

Algorithm 3: Hybrid Generation Strategy using GAN and Diffusion Models

Input: Generator G, Diffusion model D, Configuration parameters C, Number of samples N
Output: GAN samples S_{GAN}, Diffusion samples S_{Diff}, Hybrid samples S_{Hybrid}

Step 1: Initialize
Set device $\leftarrow C.device$
Set latent dimension $z \leftarrow C.gan\_latent$
Set sequence length $L \leftarrow C.seq\_len$

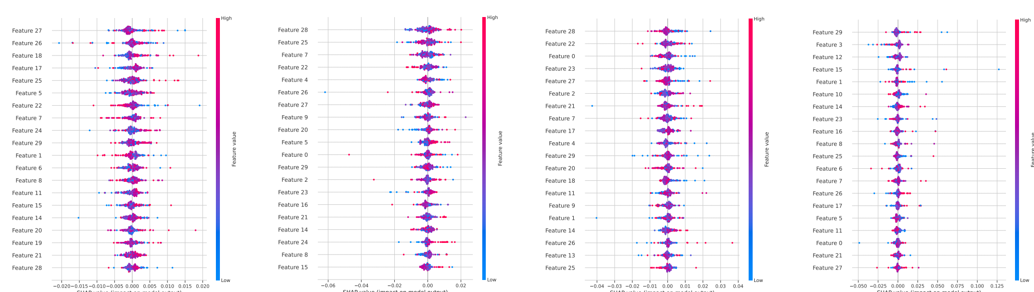
Step 2: Generate GAN samples
Draw latent vectors $Z \sim \mathcal{N}(0, I)$ of shape (N, z)
$S_{GAN} \leftarrow G(Z)$ of shape (N, L)

Step 3: Generate Diffusion samples
Initialize $X \sim \mathcal{N}(0, I)$ of shape (N, L)
for $t = T, T-1, \ldots, 1$ **do**
  $Pred \leftarrow D(X)$
  Update $X \leftarrow X - (Pred \times (0.1 + 0.9 \cdot \frac{t}{T}))$
$S_{Diff} \leftarrow X$

Step 4: Hybrid Ensemble
$S_{Hybrid} \leftarrow 0.5 \cdot S_{GAN} + 0.5 \cdot S_{Diff}$

Return $S_{GAN}, S_{Diff}, S_{Hybrid}$

the hybrid model achieves the lowest error metrics, including an RMSE of 0.043558 for BTC-USD, highlighting its superiority for scenario generation; however, further validation under different market regimes is warranted. The performance results in Table 2 show apparent differences across the three generative approaches. GANs achieve the strongest predictive accuracy, provoking lower RMSE and MAE values than diffusion models across most assets, consistent with their ability to model short-term volatility patterns. Diffusion models, while less accurate in point prediction, demonstrate greater stability in sequence generation, reflecting their robustness in capturing complex temporal structures. The hybrid GAN–Diffusion model provides the most balanced performance, reducing overfitting and improving generalization, particularly for ETH-USD and XRP-USD. These outcomes align with recent findings that adversarial models are effective at short-horizon dynamics, whereas diffusion models offer resilience to noise in longer-horizon forecasting [24]. When compared with prior work using LSTM and transformer-based architectures for cryptocurrency volatility modeling [25–27], the hybrid framework exhibits greater adaptability across heterogeneous assets, reinforcing its suitability for scenario generation and stress-testing applications in financial risk management. Figure 1 shows that SHAP value patterns differ significantly across assets. BTC-USD and LTC-USD display strong positive contributions concentrated in recent features, indicating high sensitivity to short-term returns, while ETH-USD shows weaker and more diffuse influences. XRP-USD exhibits a more balanced distribution of features. These patterns align with the performance metrics (e.g., BTC-USD Test RMSE = 0.0377), confirming that assets with stronger feature concentration tend to yield more stable predictions. Overall, Figure 1 highlights distinct feature relevance structures across cryptocurrencies, supporting model interpretability for stress testing and risk assessment.



(a) BTC-USD (b) ETH-USD (c) XRP-USD (d) LTC-USD

Figure 1: Distribution of SHAP values for BTC-USD, ETH-USD, XRP-USD, and LTC-USD

Table 2: Performance metrics for BTC-USD, ETH-USD, XRP-USD, and LTC-USD using GAN, Diffusion, Hybrid models, and Surrogate (Train/Test).

		Surrogate (XGBoost)		**GAN**		**Diffusion**		**Hybrid (GAN+Diff)**	
Cryptocurrency	**Metric**	**Train**	**Test**	**Sequence**	**Prediction**	**Sequence**	**Prediction**	**Sequence**	**Prediction**
BTC-USD	RMSE	0.015373	0.037694	0.049264	0.023051	0.171186	0.058377	0.091926	0.043558
	MAE	0.011094	0.026042	0.037962	0.015781	0.133483	0.046367	0.071809	0.034177
	MAPE	2.323200	4.619500	12.44110	10.87470	50.45780	67.02550	26.34820	30.68830
ETH-USD	RMSE	0.016348	0.049089	0.079668	0.055723	0.112913	0.065895	0.083058	0.056404
	MAE	0.012013	0.034202	0.061566	0.040023	0.087618	0.050910	0.064426	0.041239
	MAPE	2.177800	2.664500	5.513700	3.695700	9.927900	9.591700	5.818400	2.793100
XRP-USD	RMSE	0.018848	0.062375	0.123112	0.106666	0.134861	0.119866	0.122966	0.109765
	MAE	0.013359	0.041351	0.077448	0.056461	0.086904	0.067185	0.075008	0.059032
	MAPE	2.191500	3.869000	4.606900	1.543800	5.997300	2.561300	3.862900	2.085700
LTC-USD	RMSE	0.021337	0.062403	0.058310	0.024301	0.094803	0.053499	0.061319	0.031336
	MAE	0.015553	0.038015	0.045331	0.017892	0.070398	0.034247	0.046316	0.021362
	MAPE	1.975500	3.163200	7.120400	8.696900	12.41060	21.93720	7.337200	12.43930

The LIME analysis reveals that GAN models construct variable and sometimes unstable local feature contributions, whereas Diffusion models frequently produce adverse or inconsistent effects. In contrast, the Hybrid model causes smoother and predominantly positive contributions across all assets, reflecting better alignment with underlying market dynamics. These patterns are consistent with the lower RMSE values observed for the Hybrid model (e.g., 0.03134 for LTC-USD, 0.056404 for ETH-USD, 0.043558 for BTC-USD), confirming its superior stability, interpretability, and suitability for scenario generation in volatile cryptocurrency markets.

4 Conclusion

The study demonstrates that the integration of GAN and Diffusion-based generative models, particularly the Hybrid approach, significantly enhances scenario generation and stress testing for cryptocurrency markets, as evidenced by improved ACF alignments, SHAP-based feature importance, and LIME value distributions for BTC-USD, ETH-USD, XRP-USD, and LTC-USD. The Hybrid model consistently outperforms standalone GAN and Diffusion models, achieving lower RMSE and MAPE values (e.g., 0.0436 and 30.6883 for BTC-USD Hybrid Prediction), and provides robust local explainability, making it a promising tool for financial risk assessment [28]. However, limitations persist, including the models' sensitivity to hyperparameter tuning, potential overfitting to specific market conditions observed in 2023–2025 data, and challenges in capturing extreme volatility events, as indicated by higher MAPE in Diffusion predictions (e.g., 67.0255 for BTC-USD). Future work should focus on developing adaptive architectures to handle dynamic market shifts, incorporating real-time data streams, and validating models across diverse economic scenarios to enhance generalizability and robustness in stress-testing applications.

References

[1] Z. Wang and C. Ventre, "A financial time series denoiser based on diffusion models," in *Proceedings of the 5th ACM International Conference on AI in Finance*, pp. 72–80, 2024.

[2] T. Takahashi and T. Mizuno, "Generation of synthetic financial time series by diffusion models," *Quantitative Finance*, pp. 1–10, 2025.

[3] M. Villaizán-Vallelado, M. Salvatori, C. Segura, and I. Arapakis, "Diffusion models for tabular data imputation and synthetic data generation," *ACM Transactions on Knowledge Discovery from Data*, vol. 19, no. 6, pp. 1–32, 2025.

[4] S. Flaig and G. Junike, "Scenario generation for market risk models using generative neural networks," *Risks*, vol. 10, no. 11, p. 199, 2022.

[5] H. Oukhouya, F. E. Lfaze, R. Guerbaz, K. Belkhoutout, A. Lmakri, M. Fihri, L. Krishnasamy, and A. El Afia, "Forecasting stock markets in the mena region: Arimax and ensemble machine learning models with shap interpretability," *Quality & Quantity*, pp. 1–32, 2025.

[6] F. Abid, S. Bahloul, and M. Mroua, "Financial development and economic growth in mena countries," *Journal of Policy Modeling*, vol. 38, no. 6, pp. 1099–1117, 2016.

[7] M. Manian and N. Rastin, "Bitcoin price forecasting with generative adversarial networks," in *2025 29th International Computer Conference, Computer Society of Iran (CSICC)*, pp. 1–7, IEEE, 2025.

[8] G. Kim, S.-Y. Choi, and Y. Kim, "A diffusion-based generative model for financial time series via geometric brownian motion," *arXiv preprint arXiv:2507.19003*, 2025.

[9] H. Huang, M. Chen, and X. Qiao, "Generative learning for financial time series with irregular and scale-invariant patterns," in *The Twelfth International Conference on Learning Representations*, 2024.

[10] Y. Tanaka, R. Hashimoto, T. Takayanagi, Z. Piao, Y. Murayama, and K. Izumi, "Cofindiff: Controllable financial diffusion model for time series generation," *arXiv preprint arXiv:2503.04164*, 2025.

[11] Z. Wang, H. Zheng, P. He, W. Chen, and M. Zhou, "Diffusion-gan: Training gans with diffusion," *arXiv preprint arXiv:2206.02262*, 2022.

[12] C. Bekkaye, H. Oukhouya, T. Zari, R. Guerbaz, and H. El Bouanani, "Generative hybrid models for fraud detection in auto insurance with a comparative analysis of vae, gan, and diffusion approaches," *Discover Artificial Intelligence*, vol. 5, no. 1, p. 313, 2025.

[13] S. Kwon and Y. Lee, "Can gans learn the stylized facts of financial time series?," in *Proceedings of the 5th ACM International Conference on AI in Finance*, pp. 126–133, 2024.

[14] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.

[15] A. M. Antonopoulos and D. A. Harding, *Mastering Bitcoin: Programming the open blockchain*. " O'Reilly Media, Inc.", 2023.

[16] Y. Liu, N. Naktnasukanjn, A. Tamprasirt, and T. Rattanadamrongaksorn, "Comparison of the asymmetric relationship between bitcoin and gold, crude oil, and the us dollar before and after the covid-19 outbreak," *Journal of Risk and Financial Management*, vol. 16, no. 10, p. 455, 2023.

[17] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[18] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[19] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," in *International conference on machine learning*, pp. 2256–2265, pmlr, 2015.

[20] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[21] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," *arXiv preprint arXiv:2011.13456*, 2020.

[22] F. Eckerli and J. Osterrieder, "Generative adversarial networks in finance: an overview," *arXiv preprint arXiv:2106.06364*, 2021.

[23] L. Berti, B. Prenkaj, and P. Velardi, "Trades: Generating realistic market simulations with diffusion models," *arXiv preprint arXiv:2502.07071*, 2025.

[24] E. Brophy, Z. Wang, Q. She, and T. Ward, "Generative adversarial networks in time series: A systematic literature review," *ACM Computing Surveys*, vol. 55, no. 10, pp. 1–31, 2023.

[25] J. P. Fleischer, G. von Laszewski, C. Theran, and Y. J. Parra Bautista, "Time series analysis of cryptocurrency prices using long short-term memory," *Algorithms*, vol. 15, no. 7, p. 230, 2022.

[26] T. Kehinde, O. J. Adedokun, A. Joseph, K. M. Kabirat, H. A. Akano, and O. A. Olanrewaju, "Helformer: an attention-based deep learning model for cryptocurrency price forecasting," *Journal of Big Data*, vol. 12, no. 1, p. 81, 2025.

[27] A. Yaakoub, H. Oukhouya, M. Elhia, T. Zari, and R. Guerbaz, "Deep learning for financial time series: Does lstm outperform arima and svr in international stock market predictions?," *Statistics, Optimization & Information Computing*, 2025.

[28] H. Oukhouya, A. El Rhiouane, F. E. Lfaze, T. Zari, R. Guerbaz, K. Belkhoutout, and A. Moussi, "Unravelling complex financial fraud patterns in big data: a case study utilizing explainable ai in industry goods purchasing," *International Journal of Computers and Applications*, pp. 1–20, 2025.