Final Report: Forecasting Bitcoin Transaction Fees

Partner: Trilemma Foundation

Jenny (Yuci) Zhang, Tengwei Wang, Ximin Xu, Yajing Liu

2025-06-25

Table of contents

1 Executive Summary

Bitcoin transaction fees are highly volatile and event-driven, with annual spending exceeding billions and single-day spikes reaching over $78 million (Harper 2024). Most existing tools provide only short-horizon, heuristic-based estimates for the next few blocks (<1 hour), offering limited foresight for users aiming to optimize transaction cost or timing. This project addresses that gap by forecasting fee volatility up to 24 hours ahead, with the primary goal of identifying high-volatility periods in the fastestFee tier and a secondary goal of estimating fee magnitude within those windows.

Multiple modeling approaches were explored, including classical time series methods, tree-based models, and deep learning architectures. Traditional models captured seasonality but failed to anticipate sharp spikes. The Temporal Fusion Transformer (TFT) demonstrated the best performance, capturing complex dependencies and improving both RMSE and a custom volatility-sensitive loss by 25–35%. The final product is a modular forecasting system that combines narrative-driven notebooks with executable model pipelines and open-source infrastructure for ongoing development. While limitations remain—such as limited historical data, reliance on reactive features, and the computational cost of deep learning— the system provides a practical foundation for long-horizon fee forecasting and contributes to advancing infrastructure in the Bitcoin ecosystem.

2 Introduction

In the Bitcoin network, transaction fees fluctuate sharply due to congestion, shifting incentives, and user behavior. These spikes are driven by irregular, event-based shocks — such as NFT inscription surges, exchange batching, or market volatility — rather than recurring patterns (Harper 2024). Fee data is marked by abrupt jumps and heavily influenced by off-chain events, making short-term fee selection a major challenge for users and infrastructure providers (Wang et al. 2025).

Existing tools like Bitcoin Core's estimatesmartfee (Bitcoin Core Docs n.d.) and Mempool.space offer short-term, reactive guidance based on recent block data (Bitcoin Explorer n.d.), typically covering only the next 1–6 blocks. These methods are opaque, insensitive to external drivers of volatility, and provide no insight beyond the immediate horizon. Most public tools and research treat fee prediction as a static regression problem, overlooking the timing and structure of volatility — and leaving a critical planning gap unaddressed (Li et al. 2020).

2.1 Subsection (Use Two Hashes)

Our project addresses this gap by reframing fee prediction as a volatility-first, time-sensitive forecasting problem. For longer horizons, the timing and shape of fee spikes are often more actionable than precise point estimates. To capture these dynamics, we evaluate a diverse set of models: Holt-Winters Exponential Smoothing (HWES) (Holt 1957; Winters 1960), SARIMA (Box et al. 2015), XGBoost (Chen and Guestrin 2016), Prophet (Taylor and Letham 2018), DeepAR (Salinas, Flunkert, and Gasthaus 2017), and the Temporal Fusion Transformer (TFT) (Lim et al. 2019). These represent a progression from classical time series models to tree-based regressors and deep learning architectures. We assess them using both standard metrics (e.g., RMSE, MAPE) and a custom composite loss designed to penalize deviation from spike shape, timing, and volatility.

We find that TFT significantly outperforms other models, improving key evaluation metrics by 25–35% over baseline approaches. This demonstrates its ability to model non-periodic fee behavior and support forward-looking decision-making. The remainder of this report is organized as follows: Section III outlines the data science techniques used, including preprocessing, model selection, and evaluation strategy. Section IV presents the data product and results, detailing model performance, intended use, and the system's extensibility. Section V concludes with key takeaways, limitations, and recommendations for future development.

3 Data Science Techniques

3.1 Subsection (Use Two Hashes)

3.2 Subsection (Use Two Hashes)

3.3 Subsection (Use Two Hashes)

3.4 Subsection (Use Two Hashes)

4 Data Product and Results

4.1 Overview of the Data Product

This data product directly supports Trilemma Capital's mission of serving industry talent and advancing Bitcoin infrastructure through data science, both educational and technical. Its design intentionally tailors to three core audiences. General users and institutions can rely on the 24-hour forecasts to plan transactions and reduce fee costs. Learners and educators receive a transparent, step-by-step walkthrough of Bitcoin-fee forecasting and time-series methodology.

Industry experts and partners see infrastructure-grade modeling practice embodied in a modular pipeline and well-documented repository. The product is purposefully modular. Jupyter notebooks guide users through EDA, modeling decisions, and final TFT results. Python scripts implement a structured pipeline for reproducible experiments and easy re-training on new data. Finally, the open-source GitHub repository—with clear documentation—enables collaboration, scalability, and long-term extensibility.

4.2 Results

5 Conclusion and Recommendations

This project set out to address the challenge of forecasting Bitcoin transaction fees—a notoriously volatile and difficult-to-predict signal shaped by network congestion, user incentives, and unpredictable events. Recognizing the limitations of existing tools that offer only near-term, reactive guidance, we reframed the problem as one of volatility forecasting over a 24-hour horizon. Through extensive experimentation across statistical models, tree-based learners, and deep sequence models, we found that modern deep learning architectures—particularly the Temporal Fusion Transformer—are well-suited to capturing irregular spike patterns. The resulting system moves beyond average-point estimation to provide actionable foresight, helping users, wallets, and infrastructure providers make more informed decisions around transaction timing.

Despite these gains, several constraints remain. The dataset used spans just two months, limiting exposure to rare but impactful phenomena. Many features that help explain fee behavior—such as mempool congestion or block composition—are available only with a lag, reducing our ability to predict true leading signals. Furthermore, while deep models like TFT are powerful, they require regular retraining and high computational resources, which may constrain deployment or accessibility. Our custom loss function, though aligned with business needs, remains static and could be better tuned to balance volatility sensitivity and accuracy.

Looking forward, future work could explore alternate data sources and modeling strategies that address these constraints. First, longer historical coverage or the inclusion of off-chain indicators may improve predictive range. Second, deeper exploration and adaptive tuning of the custom loss function may further align model behavior with user pain points around volatility and timing. Third, future approaches might benefit from hybrid pipelines—where intermediate signals (e.g., mempool congestion or block inclusion rates) are predicted first and then fed into the final fee forecasting model. Finally, probabilistic forecasting and uncertainty quantification could add value once magnitude prediction improves.

6 Appendix

6.1 Terminology

Term	Definition
Bitcoin	Unit of currency is called ”bitcoin” with a small b, and system is called ”Bitcoin,” with a capital B. ”bitcoin” is a virtual currency (cryptocurrency) designed to act as money and a form of payment outside the control of any one person, group, or entity (i.e. decentralized).
Bitcoin Address	“1DSrfJdB2AnWaFNgSbv3MZC2m74996JafV” An encoded base58-check version of a public key 160-bit hash consists of a string of letters and numbers. Think of it analogous to an email address when sending someone an email.
Blockchain	A decentralized digital ledger that records transactions across a network of computers, making it transparent, immutable, and resistant to tampering. Technology used by Bitcoin.
Fees	The sender of a transaction often includes a fee to the network for processing the requested transaction. Most transactions require a minimum fee of 0.5 mBTC (millibitcoin) = 0.0005 BTC. Typical unit measurement in satoshi/bytes.
Hash	A function that converts an input of letters and numbers into an encrypted output of a fixed length. The hash is irreversible, meaning it cannot be decrypted back to the original input. Hashes are used in Bitcoin to create blocks and verify transactions.
Mempool	The bitcoin Mempool (memory pool) is a collection of all transaction data in a block that have been verified by Bitcoin nodes, but are not yet confirmed.
Mining / Miner	A process/network node that finds valid proof of work for new blocks, by repeated hashing.
Node	Refers to blockchain stakeholders and their devices that keep a copy of the distributed ledger and serve as communication points within the network. Major purpose is to verify the validity of the transactions within a particular blockchain.
Proof-of-Work	A piece of data that requires significant computation to find; In bitcoin, miners must find a numeric solution to the SHA256 algorithm that meets a network-wide target, the difficulty target.

Satoshi	The smallest denomination of bitcoin that can be recorded on the blockchain. 1 Bitcoin is equivalent to 100 million satoshis, named after the creator of Bitcoin, Satoshi Nakamoto.

Table 1: Key Terms and Definitions in Bitcoin and Blockchain (Alphabetically Ordered)

References

Bitcoin Core Docs. n.d. "Estimatesmartfee." https://developer.bitcoin.org/reference/rpc/estimatesmartfee.html.

Bitcoin Explorer. n.d. "Mempool.space." https://mempool.space/.

Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. 5th ed. Wiley.

Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–94. https://doi.org/10.1145/2939672.2939785.

Harper, Colin. 2024. "Bitcoin Transaction Fees Hit Record Levels After Halving — Here's Why." https://www.forbes.com/sites/colinharper/2024/04/22/bitcoin-transaction-fees-hit-record-levels-after-halving---heres-why/.

Holt, Charles C. 1957. "Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages." *International Journal of Forecasting* 20 (1): 5–22. https://doi.org/10.1016/0169-2070(94)00023-X.

Li, Xiaoqi, Peng Jiang, Ting Chen, Xiapu Luo, and Qiaoyan Wen. 2020. "A Survey on the Security of Blockchain Systems." *Future Generation Computer Systems* 107: 841–53. https://doi.org/10.1016/j.future.2017.08.020.

Lim, Bryan, Sercan O. Arik, Nicolas Loeff, and Tomas Pfister. 2019. "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting." *arXiv Preprint arXiv:1912.09363*. https://arxiv.org/abs/1912.09363.

Salinas, David, Valentin Flunkert, and Jan Gasthaus. 2017. "DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks." *arXiv Preprint arXiv:1704.04110*. https://arxiv.org/abs/1704.04110.

Taylor, Sean J., and Benjamin Letham. 2018. "Forecasting at Scale." *The American Statistician* 72 (1): 37–45. https://doi.org/10.1080/00031305.2017.1380080.

Wang, Minxing, Pavel Braslavski, Vyacheslav Manevich, and Dmitry Ignatov. 2025. "Bitcoin Ordinals: Bitcoin Price and Transaction Fee Rate Predictions." *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3541302.

Winters, Peter R. 1960. "Forecasting Sales by Exponentially Weighted Moving Averages." *Management Science* 6 (3): 324–42. https://doi.org/10.1287/mnsc.6.3.324.