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

1.1 Subsection (Use Two Hashes)

2 Introduction

2.1 Subsection (Use Two Hashes)

2.2 Target Objectives & Data Motivation

To make this problem computationally tractable, we refined the initial partner question—"How can we forecast Bitcoin transaction fees?"—into tangible scientific objectives that could be addressed using data science techniques. Through exploratory analysis and iterative scoping, we ultimately framed the task as a short-term time series forecasting problem, with a primary emphasis on detecting periods of high volatility and sudden fee spikes.

Our first objective was to identify the timing of upcoming fee volatility within a 24-hour window. From the partner's perspective, the ability to foresee periods of elevated transaction cost—regardless of precise magnitude—provides actionable value. For users, it can inform decisions on whether to delay a transaction; for wallets and exchanges, it can guide cost optimization and fee recommendation strategies. Thus, this volatility-awareness framing aligns more closely with real-world needs than a pure regression approach.

A secondary objective was to forecast the magnitude of the fastest fee during volatile periods. While precise fee prediction remains challenging due to the irregular and event-driven nature of spikes, estimating directional movement or high-fee ranges can still help downstream systems adapt proactively. This reflects the broader challenge in forecasting highly volatile series, where traditional models often fall short in capturing non-repeating patterns triggered by external events (Taylor and Letham 2018).

To support these objectives, we used a dataset consisting of 5-minute snapshots from the Bitcoin mempool, covering a two-month period in early 2025. Each snapshot captures a rich set of features—including mempool congestion, block metrics, mining difficulty, and BTC price signals—yielding over 60 time-varying variables that reflect real-time network dynamics. Prior work has shown that mempool congestion and transaction characteristics are key determinants of Bitcoin transaction fees (Fan and Liu 2020). Based on the exploratory data analysis, these features exhibited strong short-term dependencies, daily cycles, and non-repeatable spike events—further justifying the choice to focus on volatility modeling.

Overall, these refined objectives reflect a compromise between practical forecasting needs and technical feasibility, allowing us to frame a modeling strategy that is both scientifically meaningful and impactful to the capstone partner's use case.

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

5.1 Subsection (Use Two Hashes)

5.2 Subsection (Use Two Hashes)

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

Fan, Ying, and Xiaotong Liu. 2020. "The Determinants of Bitcoin Transaction Fees: Evidence from the Mempool." *Finance Research Letters* 35: 101289. https://doi.org/10.1016/j.frl.2019.101289.

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.