Proposal Report: Forecasting Bitcoin Transaction Fees

Partner: Trilemma Foundation

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

2025-05-09

Table of contents

1 Executive Summary

This project aims to develop a machine learning-based tool to forecast Bitcoin transaction fees over a 24-hour horizon. Existing tools are limited to short-term forecasts (1-6 blocks), which are insufficient for users planning batch transactions or optimizing timing. By incorporating confidence intervals, our system helps users assess both expected fees and their reliability. The tool will be deployed using AWS Lambda for real-time predictions based on mempool data, including transaction volume, fee rates, and network conditions. Baseline models such as ARIMA and XGBoost will be compared with advanced models like Prophet and DeepAR to ensure both accuracy and robustness.

2 Introduction

Bitcoin transaction fees are determined by a market-based mechanism where users compete for limited block space. This causes fees to fluctuate significantly within short timeframes, making it difficult for users to plan batch transactions or optimize timing. Most existing tools offer only short-term forecasts—typically within the next 10 to 60 minutes—which are insufficient for users who need to make decisions over longer time horizons. In addition, fee-related data is often undocumented or inconsistently structured, adding complexity to model development.

To address this gap, our project proposes a machine learning-based tool that forecasts Bitcoin transaction fees up to 24 hours in advance. By leveraging real-time mempool data—including transaction volume, fee rates, and network conditions—we aim to build a robust and accurate forecasting system. This system not only predicts the expected fees but also provides confidence intervals to help users assess the reliability of the predictions. The tool will be deployed via AWS Lambda to support scalable, real-time access. Our approach involves benchmarking baseline models like ARIMA and XGBoost against more advanced models such as Prophet and DeepAR to ensure high predictive performance under volatile network conditions. To support this, we anticipate the need for scalable training environments—especially for high-capacity models like TFT—and may explore partnerships or cloud-based solutions to extend the training horizon.

3 Data Overview

4 Data Science Techniques

5 Workflow & Timeline

6 References