Analyzing NFT Transactions and Forecasting Market Trends Using Smart Contract Data and LLMs

Yuan Jiang (yj2848), **Junfeng Zou (jz3850)**, **Rosie Wang (rw3085)**, **Huayuan Jiang (hj2683)**, **Xiaohang Ding (xd2330)**

ABSTRACT

The project involves mining OpenSea's Seaport protocol trading activity data using GraphQL queries as the base and subsequently organizing the data into well-structured, time-stamped spreadsheets. Exploiting LLMs, including GPT and DeepSeek, we are to obtain interpretable abstracts, point at anomalies, and predict near-term market trends instead. Combining blockchain data extraction and AI-driven analysis, the project develops a scalable infrastructure for translating unstructured smart contract data into relevant information on which traders, analysts, and developers in the NFT ecosystem can make decisions.

KEYWORDS

NFT Transactions, Smart Contracts, The Graph Protocol, Large Language Models (LLMs), Market Trend Prediction

1 INTRODUCTION

The newly launched NFTs for the very first time redefine the notion of digital ownership. The new phenomenon upholds transparency of origin, proves the ownership, and allows the digital assets to be scarce as well as fungible. As this is the case, it will be hard for them to analyze trading patterns, notice anomalies, and make price forecasts because the industry tools tend to not be solid enough. On the platform, users normally receive patchy and sometimes untimely information due to the sluggishness of on-chain activity. Consequently, the traders and developers will not have relevant data to make their decisions.

Such difficulties are resolved in this project by proposing a novel technique that incorporates blockchain data extraction and LLMs as analysis tools to interpret NFT market behavior. We concentrate on the OpenSea Seaport Protocol, a highly developed marketplace for non-fungible tokens (NFTs), and filter out transaction-level data by indexing the OrderFulfilled event with The Graph Protocol. Subsequently, a pipeline, a custom Python script, processes this on-chain raw data to a structured CSV dataset, which constitutes timestamp normalization, address standardization, and ETH price conversion.

Co-morbid factor input for statistical modeling also covers LLM-based fact interpretation. In this case, we conduct linear regression on the collected data to correlate with the price fluctuations while also initiating advanced LLMs (like GPT4o and DeepSeek) to summarize the transaction patterns, spot the anomalies, and produce short-term forecasts. Conclusively, the project presents a way to combine the figuration of smart contract data together with reasoning features of language models and define an interpretable and scalable method to NFT market analysis.

However, our innovation is developing a decentralized data extraction to the AI-empowered trend analysis, making it the basic pipeline that turns the raw blockchain events data into NFT market insights. The whole code as well as the dataset are archived on our GitHub, thus providing a feasible yet broad groundwork for NFT traders, analysts, and developers.

2 METHODOLOGY

This part explains how to create the entire process that is the main focus of this project. The methodology is divided into three stages: the first stage involves selecting and deploying smart contracts, after that we proceed with on-chain data extraction and API design, and lastly we apply large language model (LLM) for market trend analysis.

2.1 Smart Contract Selection and Deployment

The foundation of this project is the **OpenSea Seaport Protocol** [1], a widely adopted smart contract that facilitates decentralized NFT trading on the Ethereum blockchain. The contract is deployed at the address:

`0x00000000006c3852cbEf3e08E8dF289169EdE581`

This protocol lays out the definition of NFT trades, such as creation, execution, and `cancellation`, which use public event logs to provide this governance mechanism with transparency and immutability. The transaction can be done by two primary functions, namely `fulfillOrder` and `cancelOrder`.

The `fulfillOrder` function is the main tool for carrying out a deal between the people who receive the offers (the sellers) and those who accept the offers (the buyers). First, it takes a full order description, such as item type, token address, amount, ID, and other conditions, and then it settles the asset transfer at a single point in time. This function turns on the `OrderFulfilled` event as an outcome of fulfillment, and allows downstream systems to trace the sales parameters.

The `cancelOrder` function gives a seller the possibility to nullify the offer that they previously have made before it is fulfilled. This maintains the on-chain resilience, so that the out-of-date or undesired listings can be revoked in the blockchain. These vectors together form the Seaport's core exchange engine's structural process.

A point of focus is the `OrderFulfilled` event, which has a record of all the main transaction data, including `offerer` and `fulfiller` addresses, `NFT` contract address, `token ID`, `payment token`, price (in `amount`), and the transaction timestamp. This event supplements the creation of the formalized sales records that reflect the actual buying and selling activities made on OpenSea.

At the beginning of the scheme, we meant to conduct research on the trading behavior and reap future market trend predictions, so we set up on-chain `NFT` transactions from the Seaport protocol. Along with anticipating floor price changes in the short run, we aspired to give conciseness – spotting transaction trends, rushes, and unusual movements awareness, and quarterly seasonality of the **LLMs** (*large language models*).

2.2 Data Extraction and API Design

To obtain the transaction-level data from the Seaport smart contract, we employed **The Graph Protocol** [2]. A dedicated subgraph [3] had been built for embedding the `OrderFulfilled` events based on the said contract. The process encompassed registering the data schema and event mappings, alongside the `GraphQL` interface, as follows:

- **schema.graphql**: defines an `NFTSale` entity with fields including `id`, `collection`, `tokenId`, `price`, `timestamp`, `txHash`, and `paymentToken`.
- **subgraph.yaml**: specifies the data source (Seaport contract), target network, ABI file, and event handler.
- **mapping.ts**: processes relevant event data and maps it into the schema-defined entity.

After deploying it via **The Graph Studio**, the subgraph pinpoints the NFT's real-time transaction using a `GraphQL` interface.

In our approach, we used a Python script (named `fetch_sales.py`) to collect and clean up the indexed data. This script sends `GraphQL` requests, transforms relevant fields (e.g., from `wei` to ETH), and finally exports the result to a CSV file called `nft_sales.csv`.

The dataset includes:

- `id`: Transaction hash + log index
- `collection`: NFT contract address
- `tokenId`: NFT identifier
- `price`: Sale price in ETH
- `paymentToken`: Currency used for payment
- `timestamp`: Unix timestamp
- `txHash`: Transaction hash
- `datetime`: ISO-formatted human-readable time

2.3 LLM Integration

Subjecting the structured data about the NFT sales to an LLM, such as GPT or DeepSeek, they will lead into retrieval of history trends and generation of forecasts for future trading activity. By prompting these models with specific questions – such as predicting price movements or identifying anomalous trading patterns – the project demonstrates how AI can enhance blockchain analytics. The LLMs carry out an analysis based on both quantitative forecasts and qualified knowledge as a result of which a complete 30 days' forecast can be given. The prompt to the DeepSeek is shown as below:

Please analyze the NFT transaction data in the csv file, summarize the transaction trends in the past few months, including changes in transaction volume and possible seasonal patterns, and answer the following questions:

1. What is the time range of the data? Does it influence the change in the trading frequency?

2. Can you tell that trained volume has any peaks or troughs? If so, why?

3. Check whether there are any abnormal transactions (such as high-frequency trading or price fluctuations) in the data and analyze the possible causes.

4. By using historical data, predict/forecast the trading direction for the next 30 days and clarify the causes.

3 RESULTS

To complement our statistical analysis, we submitted `nft_sales.csv` dataset to DeepSeek-V2, a large language model capable of extracting patterns, summarizing trading behaviors, and predicting market dynamics. The model was prompted with four questions regarding time range, volume trends, anomalies, and market projections.

3.1 Time Range and Trading Frequency

DeepSeek uncovered a time span of 10 months that included a range from June 18, 2024, through April 3, 2025. Activity in the markets during the first months from June to August remains low, with transactions totaling around 5 to 10 per month. There was an accelerating trend at the end of 2024, where the sales started to peak with 30 sales in December.

The biggest rise in the dataset happened in March 2025, when over 50 transactions occurred—the highest volume month in the dataset. With such a sudden rise, the data map showed an instant drop to two transactions in early April 2025 being indicative of a massive market crunch.

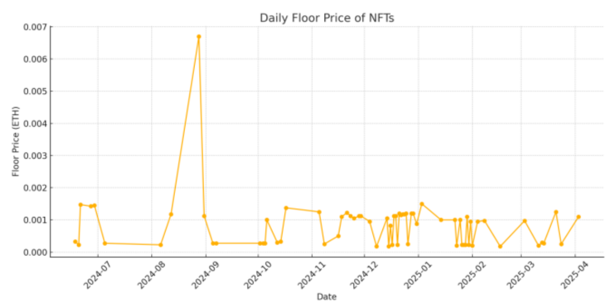

Figure 1: Monthly transaction volume trend from June 2024 to April 2025

3.2 Market Peaks and Troughs

The language model previewed in Section 4 indicated a clear upward trend for March 2025 in trading volume, probably due to speculation interest and NFT collections launches, or social media engagements. In December 2024, the peaks were observed for the second time, which could have been because of the end-of-year customer spending behavior and promotional campaigns.

Conversely, DeepSeek highlighted that trading activity in the months of June and July 2024 was notably low, which possibly occurred due to the seasonal dips or the sticky market consolidation phases during which the demand-side engagement of the buyers shrinks.

3.3 Transaction Anomalies

On the backside of the coin, a few conspicuous deviations were isolated in the data. For instance, on August 28, 2024, a transaction of 0.0067 ETH with the token from collection `0xfd2b5dc2` was recorded. This transaction was not only above but also beyond the average price bracket and could have marked the selling of an exclusive and collectible NFT.

Moreover, DeepSeek found out that on March 11, 2025, there was a little jump of trading activities, with more than ten transactions involved over a 15-minute interval. These high-frequency mixes indicate that an automated bot activity or a group of people could be purchasing items.

An instance of price changes, including ETH values ranging from 0.000175 ETH through 0.0067 ETH, were also recorded, particularly in Q1 2025.

3.4 Market Forecast for the Next 30 Days

Among the indicators of the following new trading rounds and seasonal cycles, DeepSeek accordingly predicted a forecast with modest contraction of NFT trading in April and May 2025. This would follow the large speculative growth seen in March, which was perceived as a bubble rather than a permanent uptrend.

The forecast position assumes the same downturn as seen in mid-2024 connected with observed NFT markets have a cyclic pattern corresponding with the calendar seasons and release schedules. While no new launches or catalysts are expected to boost activity, the market looks set to stabilize at lower transaction volumes for the period ahead.

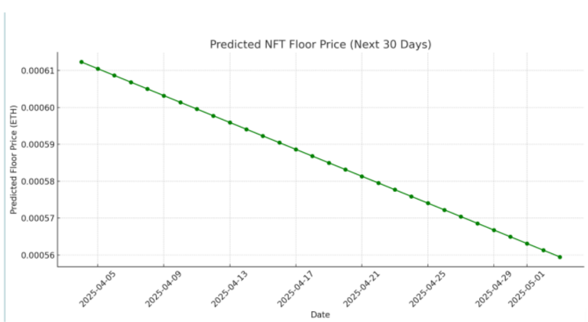

Figure 2: Predicted NFT Floor Price for Next 30 Days

4 DISCUSSION AND FUTURE WORK

Whereas the present model exemplarily integrates LLMs functioning based on decentralized information gathering, there are some more chances for its further expansion and upgrade.

4.1 Model Expansion and Comparative Analysis

In this approach, DeepSeek-V2 is the main LLM model, which guides the hints extraction. Moving forward, it is also planned to perform comparative analysis of other models (e.g., Claude and LLaMA-3) with regards to their correlation, parity, and domain capability variations in financial reasoning tasks.

4.2 Data Scope and Granularity

Currently, we have a dataset containing a total of 250 transactions during a ten-month time frame. Letting the observation timeframe go for longer, adding more data volume, or introducing more sophisticated attributes (for example, sellers' and buyers' addresses, gas fees, royalties) could help to deliver more precise forecasts and better understanding of anomalies.

4.3 Predictive Modeling and Automation

In particular, a simple linear regression was used to develop a baseline, but there is an opportunity to improve this with more advanced time-series methods, including ARIMA, Prophet, and LSTM, among others. Likewise, real-time sales data from NFT could be retrieved through automated pipelines and reportable to LLMs.

4.4 Visualization and Interface Development

A step beyond will see a user-facing dashboard built, visualizing market trends, model outputs, and red-flagged anomalies in real time. Consequently, it will play a vital role in involving the analytical insights generated through this system by the non-tech people.

5 CONCLUSION

In this project, we created an all-in-one procedure for NFT transaction data extraction, structuring, and analysis from the OpenSea Seaport Protocol. As a result of deploying a unique subgraph mechanism on The Graph, we got the `OrderFulfilled` event indexed with ease and sourced structured trading records. The processing and outputting of the data into a CSV file were done by a Python script, which allowed for both the statistical study and language model analysis.

Indeed, we find that mega-data generated out of blockchain, combined with large language models, can be translated into understandable insights into the market behavior. DeepSeek served us to find trading trends, seasonality, anomalies, and near-term forecasts. The data storage aligns intensely with the data pattern, which shows that the computer-assisted analysis is appreciated. This work fetches a manageable structure that can be further utilized as a knowledge base for Web3 market analysis, instantaneous analytics, and algorithms to deal with cryptocurrency finance.

ACKNOWLEDGMENTS

This project was supported by the ELENE6883 Introduction to Blockchain Technology course, with guidance from Professor Chong Li, Professor Chonggang Wang, and assistance from teaching assistants Yubo Zhang.

REFERENCES

[1] Etherscan. Seaport protocol smart contract. https://etherscan.io/address/0x00000000006c3852cbef3e08e8df289169ede581, 2024.

[2] The Graph. The graph protocol homepage. https://thegraph.com/, 2024.

[3] Jamie Barton. Indexing smart contract data using the graph protocol. https://dev.to/jamiescript/indexing-smart-contract-data-using-the-graph-protocol-4h95, 2022.