HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

GRADUATION THESIS

Development of a Multi-Factor Scoring Model for Evaluating Investment Potential of NFTs

Hoàng Gia Nguyên
nguyen.hg204889@sis.hust.edu.vn

Major:Data Science and Artificial Intelligence

Supervisor: Assoc. Prof. Dr. Nguyễn Bình Minh _____

Signature

Department: Computer Engineering

School: School of Information and Communications Technology

HANOI, 12/2024

ACKNOWLEDGMENT

I would like to express my heartfelt gratitude to all the teachers and professors at Hanoi University of Science and Technology, particularly those at the School of Information and Communication Technology, and more specifically, the Department of Computer Science. Throughout my studies, they have imparted invaluable lessons and provided me with a strong foundation of knowledge, enabling me to successfully complete this thesis. This foundation has been a crucial stepping stone for my future career development.

I am also deeply thankful to the teachers, mentors, and friends at A* Group for their unwavering support and assistance. Additionally, I owe special thanks to Mr Nguyen Viet Bang for helping me alot with this thesis. Your encouragement, guidance, and dedication have been incredibly meaningful to me. Your consistent support and efforts have inspired me greatly, and I am truly appreciative of everything you have done.

With profound gratitude, I extend my sincere thanks to Associate Professor Dr. Nguyen Binh Minh for his guidance throughout this thesis. Your mentorship and expertise have been instrumental in the completion of my work. I am deeply grateful for your invaluable insights and assistance.

Lastly, I would like to thank my family and friends for their continuous encouragement and support throughout my studies and the thesis-writing process.

ABSTRACT

Non-Fungible Tokens (NFTs) have revolutionized the digital economy by introducing a new class of assets that blend technology, art, and investment opportunities. However, the evaluation of NFTs for investment purposes remains a challenging task due to their diverse nature, varying market dynamics, and limited standardized methodologies. This thesis addresses the critical need for a systematic approach to assess and rank NFTs based on their investment potential.

A comprehensive scoring formula is proposed, designed to quantify the investment worth of NFTs. The formula incorporates multiple factors such as Return on Investment (ROI), trading volume, reputation scores, transaction count, and unique buyer and seller activity, ensuring a holistic evaluation of each NFT. The scoring system is calibrated to range from 0 to 1000, providing an intuitive framework for categorizing NFTs into investment tiers, from "Poor" to "Excellent."

The research further develops a robust data pipeline utilizing advanced web-scraping techniques, including Selenium, to collect real-time sales and transactional data from marketplaces. The analysis leverages ROI calculations derived from consecutive sales events, enabling investors to identify patterns and profitability trends across collections. Additionally, the thesis integrates metadata and reputation metrics from social platforms to assess the broader market sentiment and influence.

The results of the proposed methodology demonstrate its effectiveness in providing clear, data-driven insights for NFT investors and collectors. By introducing a standardized scoring framework, this thesis contributes to the growing field of NFT investment analysis, offering a tool that bridges the gap between traditional asset evaluation methods and the unique attributes of NFTs. The findings pave the way for future research into enhancing NFT evaluation techniques and exploring their applications in a rapidly evolving digital economy.

Student

(Signature and full name)

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION 1

1.1 Problem Statement 1

1.2 Background and Problems of Research 1

1.3 Research Objectives and Conceptual Framework 2

1.4 Contributions 3

1.5 Organization of Thesis 4

CHAPTER 2. LITERATURE REVIEW 6

2.1 Scope of Research 6

2.2 Related Works 7

2.2.1 Current NFT Evaluation and Scoring Methods 7

2.2.2 Traditional Asset Valuation and Scoring Models 8

2.2.3 NFT ROI Score 8

2.2.4 NFT Reputation Score 9

2.2.5 Other NFT Evaluation Methods 10

2.3 Blockchain 10

2.4 NFT 12

2.4.1 Basic Concept of NFT 12

2.4.2 NFT Key Terms 12

2.4.3 NFT Purpose and Use Cases 13

2.4.4 NFT Standards 14

2.4.5 NFT Categories 15

CHAPTER 3. METHODOLOGY 17

3.1 Overview 17

3.2 Investment Worth Score 17

3.2.1 Parameters 17

3.2.2 ROI score 18

3.2.3 Reputation Score 20

3.2.4 Proposed Formula 22

3.3 Model for Optimizer 23

3.4 DATA COLLECTION AND PROCESSING 26

3.4.1 Data Pipeline 26

3.4.2 Scraper 28

3.4.3 Calculator 31

3.4.4 Database 34

3.4.5 Labeling and Exploratory Data Analysis 41

CHAPTER 4. NUMERICAL RESULTS 48

4.1 Metrics for model's performance 48

4.2 Evaluation results 50

CHAPTER 5. CONCLUSIONS 53

5.1 Summary 53

5.2 Suggestion for Future Works 53

REFERENCE 56

LIST OF FIGURES

Figure 3.1 Gradient Descent 24

Figure 3.2 Adam Optimizer 26

Figure 3.3 Genetic Algorithm 27

Figure 3.4 Scraper pipeline 28

Figure 3.5 Calculator pipeline 32

Figure 3.6 Calculator pipeline 34

Figure 3.7 Collection Documents 35

Figure 3.8 NFT Documents 36

Figure 3.9 Sales Documents 38

Figure 3.10 EDA 43

Figure 3.11 Reputation Score EDA 44

Figure 3.12 Reputation Score EDA(remove outliers) 45

Figure 3.13 Reputation Score EDA(remove outliers) 46

Figure 4.1 Model Performance 50

Figure 4.2 Model Performance 51

LIST OF TABLES

Table 3.1 Investment Worth Score Parameters 18
Table 3.2 Required data and sources for the NFT evaluation process. . . 27
Table 3.3 API used to collect NFT data from OpenSea. 31
Table 3.4 NFT Collection Data Description 35
Table 3.5 NFT Data Description 37
Table 3.6 Sale Events Data Description 40

LIST OF ABBREVIATIONS

Abbreviation	Definition
API	Application Programming Interface
GA	Genetic Algorithm
NFT	Non-Fungible oken
ROI	Return On Investment
SGD	Stochastic Gradient Descent

CHAPTER 1. INTRODUCTION

1.1 Problem Statement

The NFT (Non-Fungible Token) market has grown exponentially in recent years, creating significant opportunities for investors and collectors. NFTs, unique digital assets tied to blockchain technology, are valued based on a wide range of factors, including rarity, utility, market trends, and reputation. However, this rapidly expanding market is characterized by a lack of standardized and reliable evaluation frameworks, making it difficult for investors to assess the actual worth of NFTs. Current pricing mechanisms are highly subjective and often influenced by external factors such as social media hype, celebrity endorsements, and speculative activities, leading to significant price volatility and investment risks. Without a systematic way to evaluate NFTs, investors face uncertainty in determining which assets hold genuine long-term value and which are overvalued or part of speculative bubbles. Market inefficiencies exacerbate these challenges, as NFTs are often priced without a clear connection to their underlying utility or potential returns. Furthermore, existing studies and methodologies on NFT valuation remain limited, leaving investors with insufficient tools to make data-driven decisions. This research seeks to address these challenges by developing an optimized formula to evaluate and rank NFTs for investment purposes. The proposed formula incorporates critical factors such as return on investment (ROI), transaction volumes, buyer and seller activity, ownership distribution, and reputation scores. By integrating these metrics into a cohesive scoring system, the formula aims to provide a comprehensive framework for assessing NFT value. This approach not only empowers investors with actionable insights but also fosters transparency and efficiency within the NFT market, reducing speculative risks and encouraging informed decision-making.

1.2 Background and Problems of Research

The rapid growth of the blockchain and cryptocurrency ecosystem has given rise to various digital assets, including NFTs (non-fungible tokens). NFTs, unique digital tokens representing ownership of a specific asset, have gained widespread popularity across multiple domains such as art, gaming, and collectibles. Despite their rapid adoption, the evaluation of NFT collections in terms of market performance, reputation, and user engagement remains a significant challenge. The absence of standardized labeling and ranking frameworks often leads to inconsistent assessments, making it difficult for investors and users to navigate the market effectively.

One of the primary problems is the lack of systematic methodologies to classify NFT collections based on their performance metrics such as transaction volumes, number of buyers, sellers, and owners. While raw numerical data is readily available, its interpretation for meaningful insights remains problematic. Additionally, the influence of social media activity on NFT reputation and valuation has been recognized but is not systematically incorporated into ranking models. Metrics like likes, retweets, and tweets mentioning collections on platforms such as Twitter can provide valuable insights into user sentiment and engagement but require advanced computational methods for analysis.

Moreover, the volatility and dynamic nature of the NFT market exacerbate these challenges. Trends and market behaviors can shift rapidly, requiring continuous updates to ranking methodologies. The lack of a unified framework for assessing NFT collections creates an uneven playing field, leaving room for manipulation or biased assessments. These issues highlight the necessity for robust and scalable solutions that integrate quantitative performance metrics with qualitative indicators like social media influence.

This research aims to address these challenges by developing a comprehensive labeling and ranking system for NFT collections. By utilizing ROI scores, quantile-based classification, and social media data, the study seeks to establish a standardized framework that enables more transparent and reliable evaluations of NFT collections. The work not only addresses the current gaps in NFT market analysis but also contributes to advancing methodologies in the broader field of blockchain-based asset evaluation.

1.3 Research Objectives and Conceptual Framework

The primary objective of this research is to develop a comprehensive reputation scoring model for NFT collections by integrating both on-chain and off-chain metrics to assess their credibility, trustworthiness, and community engagement. On-chain data such as trading volume, transaction frequency, unique buyers and sellers, and ownership distribution provide insight into the market activity and liquidity of NFT projects. Meanwhile, off-chain factors, including social media metrics like likes, retweets, and mentions, as well as community activity on platforms such as Twitter, offer a qualitative evaluation of public perception and community involvement. By combining these dimensions, this research proposes a hybrid reputation scoring algorithm that assigns weights to key metrics, balancing market activity with social credibility. The framework incorporates a validation phase to compare the calculated reputation scores against actual market performance and user sentiment,