



KDD2024
BARCELONA, SPAIN

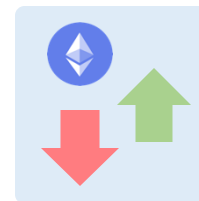
COMET: NFT Price Prediction with Wallet Profiling



NFT



COMET



Price

Tianfu Wang, Liwei Deng, Chao Wang, Jianxun Lian,
Yue Yan, Nicholas Jing Yuan*, Qi Zhang, Hui Xiong*



USTC



Microsoft



HKUST-GZ

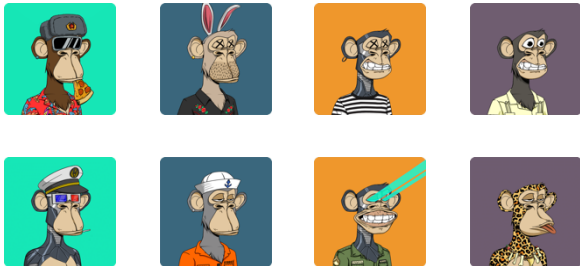
Non-fungible Token (NFT)

NFTs are unique digital assets that certify ownership and authenticity via blockchain.

- ▶ Uniqueness: Each NFT is distinct and topically associated with digital files
- ▶ Transparency: NFT Metadata and Transactions are publicly accessible

Hierarchy & Heterogeneity

Collection Consists of Tokens



Similar style,
but Different properties

[Link: Bored Ape Yacht Club](#)

Multimodal Contents

Textual Description and Visual File

NFT Collection Name & Desc.

Name	Description
Bored Ape Yacht Club	The Bored Ape Yacht Club is a collection of 10,000 ...

NFT Token File & Properties

	BACKGROUND Blue 12% Floor: 10.89 ETH	CLOTHES Toga 2% Floor: 11.6 ETH	EYES Heart 4% Floor: 11.6 ETH
	FUR Tan 6% Floor: 11.6 ETH	HAT Baby's Bonnet 2% Floor: 11.6 ETH	MOUTH Bored 23% Floor: 10.675 ETH

 NFT Collection

 NFT Token

Multi-behavior Transactions

Wallet-Blockchain & Wallet-Wallet



 Blockchain

 User Wallet

Research Objective

Existing Challenges

Ambiguous and incomplete problem definitions and evaluation method

Leverage valuable information from publicly accessible user behaviors

Identify potential price trend correlations among NFTs and between NFTs and wallets

Main Contributions of Our Work

Addressing the Gaps

Definitions and Evaluation

Mining Social Impact

Wallet Profiling with Graph

Practical Deployment

Infra. & Product Application

Work	Problem Definition			Methodology	Evaluation	
	Price-level	Target	Task Type		Dataset	Metrics
[25]	Token	Asset Value	Correlation	Correlation Analysis	1.4M Tokens	Fitting error
[35]	Token	Asset Value	Correlation	Correlation Analysis	48K Tokens	R^2
[9]	Token	Asset Value	Classification	Multimodal Learning	4.7M Tokens	Precision; Recall; F1
[19]	Token	Asset Value	Classification	LightGBM; XGBoost; etc	62K Tokens	ACC; F1
[27]	Token	Asset Value	Regression	Linear Regression	4.7M Tokens	R^2
[4]	Collection	Daily Price	Correlation	Correlation Analysis	4 Collections	/
[23]	Collection	Daily Price	Regression	NLP sentiment and MLP	19 Collections	MAE; ACC; F1
[18]	Collection	Daily Price	Regression	RNN; Linear regression	1 Collections	MSE; ACC
Ours	Collection Token	Aggregated Price Sale Price	Classification Regression	Temporal Heterogeneous Graph Learning	100 Collections 1.5M Tokens	ACC; MCC MSE; MAE

Wallet Profiling For NFT Price Prediction

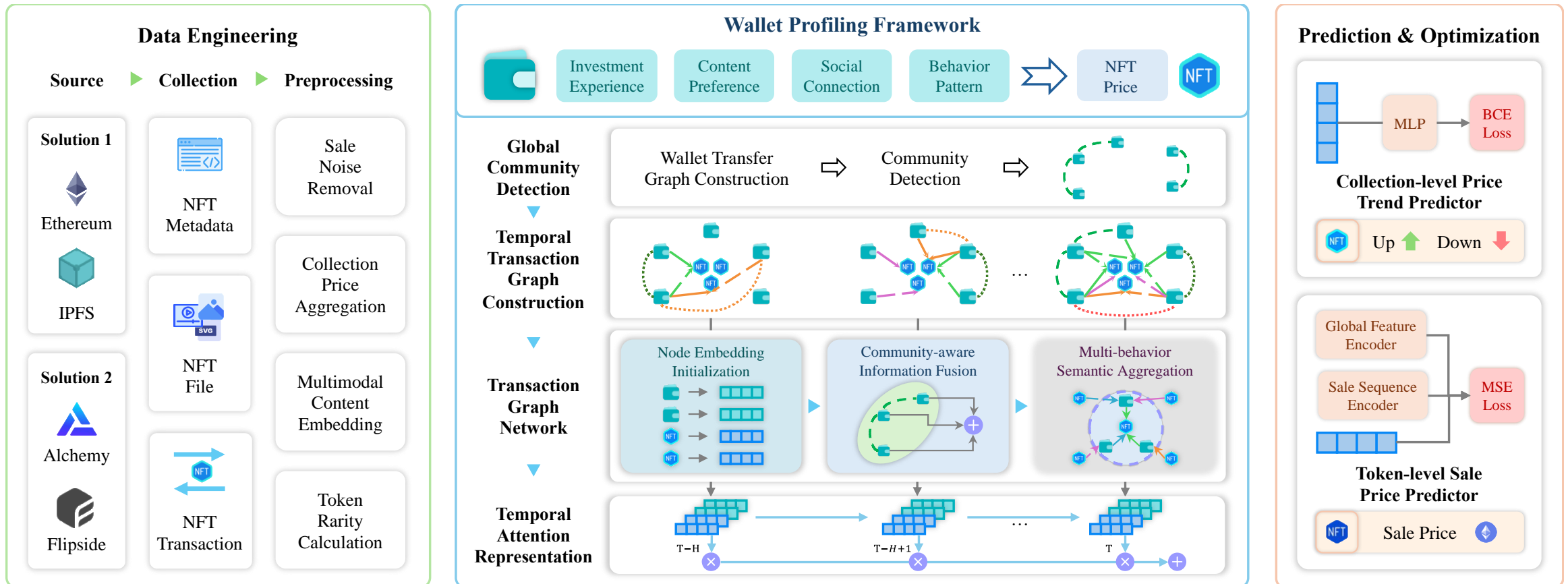
Create a comprehensive profile of a digital wallet within the NFT ecosystem.

Investment Experience
asset & profitability of wallet

Content Preference
popularity of NFT

Social Connections
latent similarity among wallets

Behavior Patterns
dynamic trading strategies



Wallet Profiling: Hierarchical Problem Definition

Both tasks at collection and token levels effectively caters to address diverse needs of users

- ▶ Manual Aggr. Price of Collection: Overall trends from a macroeconomic perspective
- ▶ Token-specific Trading Only: Detailed insights into specific decisions

Collection-level Price Trend Prediction

Allow investors to stay informed about the changing trends



Classification

Up / Down

$$\hat{y}_c = \mathcal{F}_c(\{\mathcal{X}_c^t \mid t \in [T - H, T]\})$$

Token-level Sale Price Prediction

Estimate future sale prices at which individual NFT tokens are sold



Regression

Sale Price

$$\hat{y}_u = \mathcal{F}_u(S_u; \mathcal{F}_c(\{\mathcal{X}_c^t \mid t \in [T - H, T]\}))$$

Wallet Profiling: NFT Data Engineering

Collect NFT-related data and perform pre-processing to obtain insightful data

A. Data Collection

Infra.-Supported Solution

Setup a Local Ethereum node

Scalability & efficiency & integrity

Scalable and real-time extraction

Public API-Based Solution

Alternative based on public APIs

Flexibility & Open-Source

Accessible and Convenient Tool

Data Engineering

Source ► Collection ► Preprocessing

Solution 1



Ethereum



IPFS



NFT
Metadata



NFT
File



NFT
Transaction

Sale
Noise
Removal

Collection
Price
Aggregation

Multimodal
Content
Embedding

Token
Rarity
Calculation

B. Data Preprocessing

Wash Sale Detection
Anomaly Filtering

Visual Embeddings ViT
Textual Embeddings Sentence-BERT

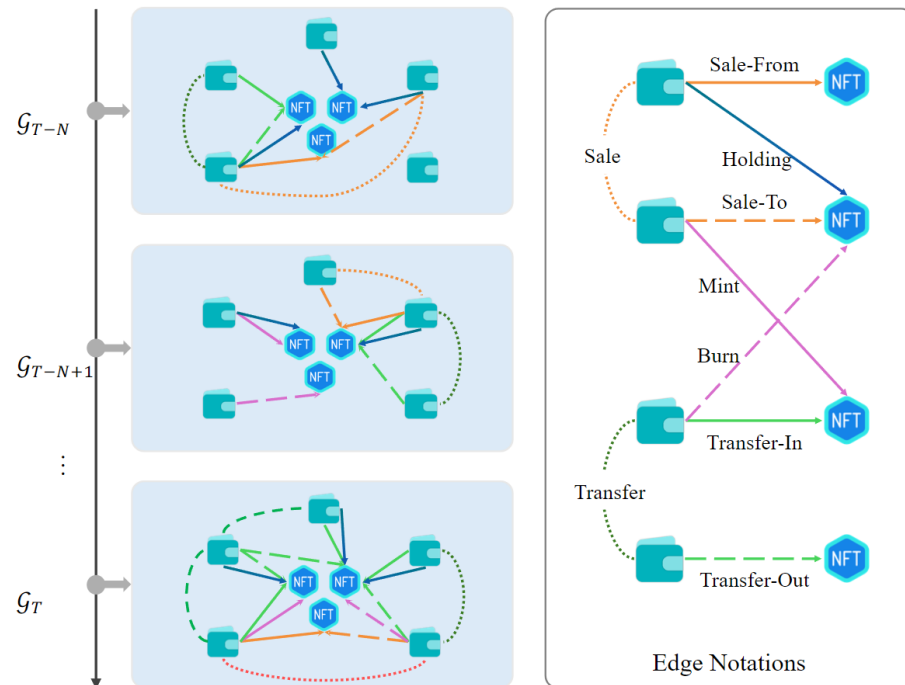
Daily Median Price
Linear Interpolation

Rarity Score Metrics
Their Influence on Price

Wallet Profiling: COMET Model

Temporal Transaction Graph Construction

- ▶ Multi-behaviours Transactions: Multifaceted interactions between wallets and collections
- ▶ Segment & Construction: A series of snapshot transaction graphs

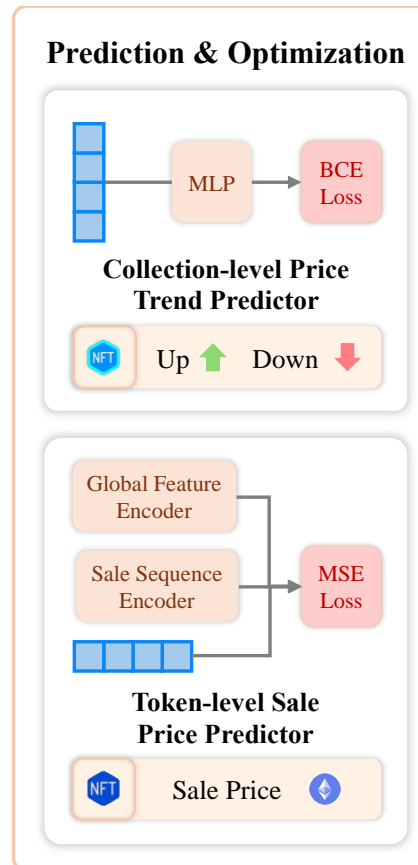
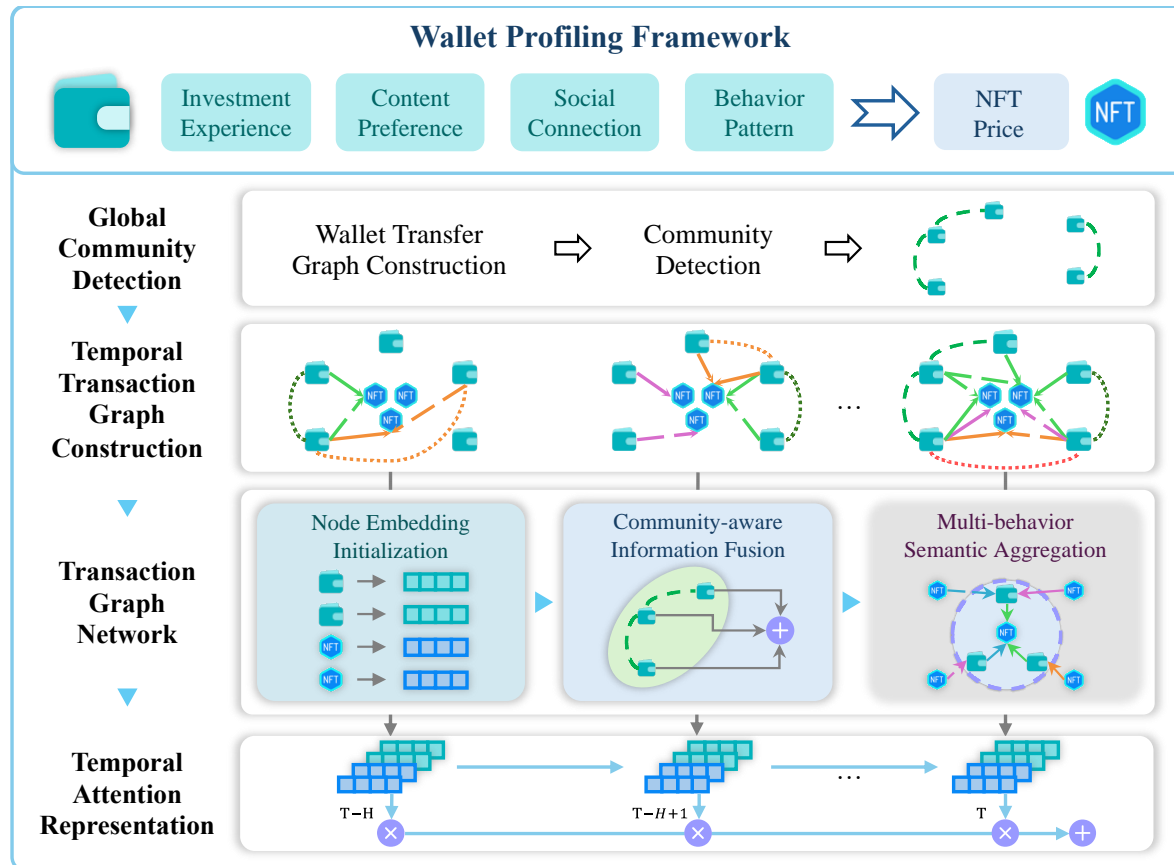


Node / Edge	Feature Type	Feature Name	Description
Collection Node	Static	Visual embedding	An embedding representing the visual content associated with the collection.
	Static	Textual embedding	An embedding representing the textual content associated with the collection.
	Static	Total count of token supply	Total number of tokens supplied in the collection.
	Dynamic	Daily price	Daily price of the collection, providing a historical price trend.
	Dynamic	Daily transaction counts	Daily counts of mints, sales, transfers, and burns of the collection.
	Dynamic	ETH-to-USD exchange rate	Daily exchange rate of Ethereum (ETH) to US Dollars (USD).
Wallet Node	Dynamic	Total sale Volume	Total sale volume of the collection over one day.
	Dynamic	Daily transaction counts	Daily counts of mints, sales, transfers, and burns of the wallet.
	Dynamic	Total holding count	Total count of NFT tokens held by the wallet.
Holding Edge	Dynamic	Total asset value	Total value of assets held in the wallet.
Sale-from Edge	Dynamic	Owned token count	Number of NFT tokens in one collection owned by the wallet.
Sale-to Edge	Dynamic	Sale Price	Price at which an NFT token was sold.
Transfer-in Edge	-	Sale Price	Price at which an NFT token was sold.
Transfer-out Edge	-	-	-
Mint Edge	-	-	-
Burn Edge	-	-	-
Sale Edge	Dynamic	-	-
Transfer Edge	-	-	-

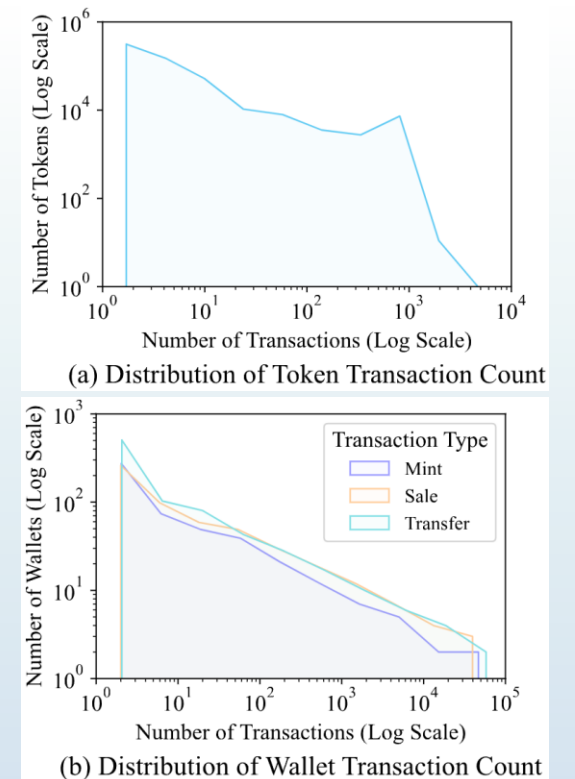
Wallet Profiling: COMET Model

Community enhanced Multi-behavior Transaction graph model

- ▶ Temporal Transaction Graph Network to capture social impact and market dynamic
- ▶ Community-aware Information Fusion Module to alleviating data sparsity issues



Issues of Data Sparsity

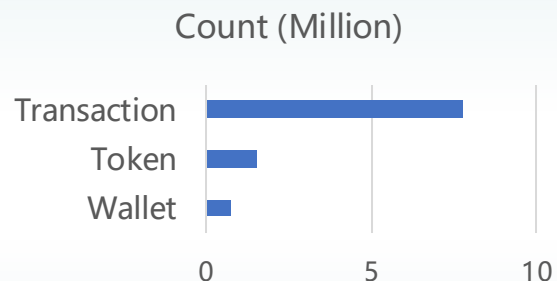


Experimental Analysis

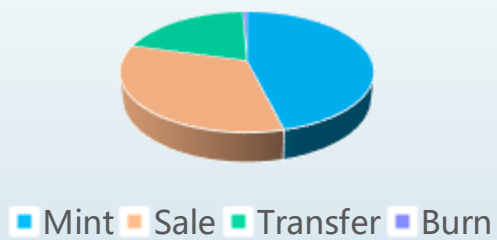
Dataset Collection

January 2021 - June 2023

Top-100 Collections



Transaction Types



A. Collection-level Task Results

Algorithm	1-step		3-step		5-step	
	ACC ↑	MCC ↑	ACC ↑	MCC ↑	ACC ↑	MCC ↑
RF	0.5627	-0.0017	0.5476	-0.003	0.5675	0.0371
SVM	0.5678	0.0537	0.5668	0.0197	0.5884	0.0419
XGBoost	0.5640	0.0665	0.5616	0.099	0.5755	0.1119
MLP	0.5513	0.0620	0.5792	0.1365	0.5808	0.1123
LSTM	0.5776	0.1372	0.5840	0.1348	0.5843	0.1237
TCN	0.5781	0.1387	0.5890	0.1386	0.5978	<u>0.1698</u>
ALSTM	<u>0.5893</u>	<u>0.1537</u>	<u>0.5977</u>	<u>0.1681</u>	0.5946	0.1329
D-Linear	0.5663	0.0451	0.5695	0.0427	0.5910	0.0680
N-BEATS	0.5835	0.1112	0.5902	0.1425	0.6029	0.0262
Informer	0.5505	-0.0173	0.5894	0.1266	0.5841	0.0262
Autoformer	0.5668	0.1063	0.5645	0.1422	<u>0.6079</u>	0.1479
COMET	0.6075	0.1861	0.6158	0.2092	0.6214	0.2128

B. Token-level Task Results

Algorithm	1-step		3-step		5-step	
	MAE ↓	MSE ↓	MAE ↓	MSE ↓	MAE ↓	MSE ↓
RF	0.4723	2.1812	0.4757	2.2164	0.4743	2.1892
SVR	0.4901	2.2098	0.4849	2.1602	0.4921	2.2178
XGBoost	0.4815	2.1524	0.4950	2.2175	0.4835	2.1604
MLP	0.4181	1.9728	0.4305	2.0321	0.4201	1.9808
LSTM	0.3689	1.775	<u>0.3638</u>	<u>1.7831</u>	0.3679	1.7836
TCN	0.3859	1.9224	0.4003	1.9808	0.3796	1.8904
ALSTM	<u>0.3631</u>	<u>1.7523</u>	0.3714	1.8054	<u>0.3601</u>	<u>1.7209</u>
COMET	0.3442	1.5825	0.3379	1.619	0.3395	1.5643
w/o CE	0.3529	1.6163	0.3424	1.6370	0.3479	1.6005
w/o TF	0.3576	1.6394	0.3561	1.6793	0.3606	1.6674

Ablation Study

Edge Ablation

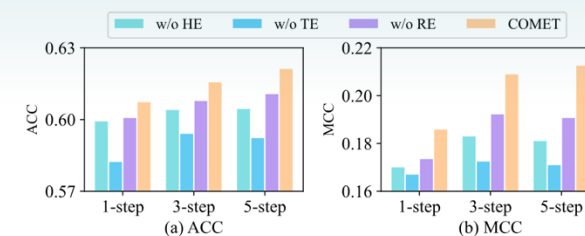


Figure 4: Results of edge ablation.

Module Ablation

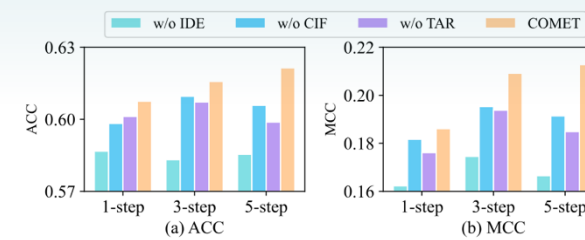


Figure 5: Results of module ablation.

Experimental Analysis

Feature Importance Analysis

Most Influential Features
historical prices
sale volumes of collections

visual > textual embeddings

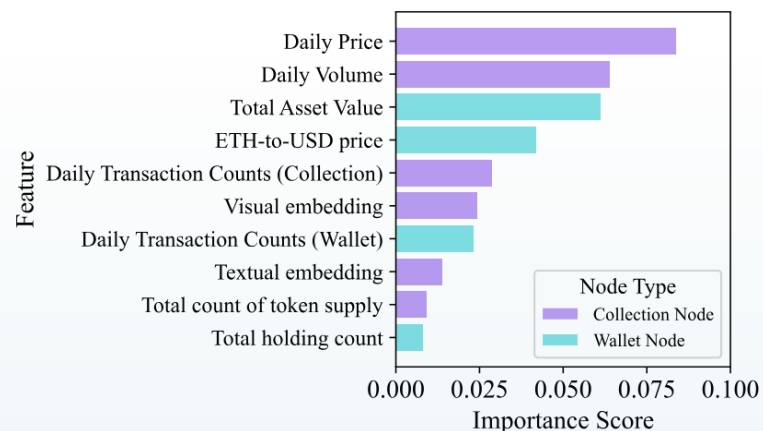
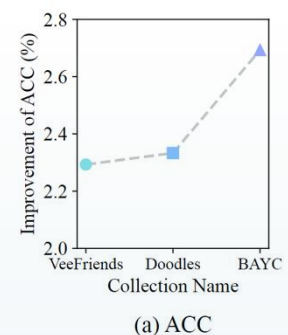


Figure 6: Top 10 feature importances.

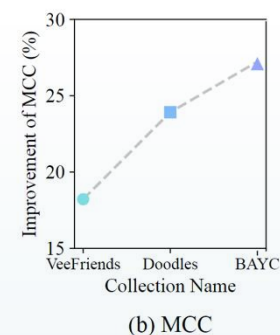
Effectiveness in Different Collections

Similar token counts
But differ in transaction counts.

Higher transaction counts
More significant improvements



(a) ACC



(b) MCC



VeeFriends
#Token: 10.3K
#Transactions: 31K



Doodles
#Token: 9.9K
#Transactions: 109K



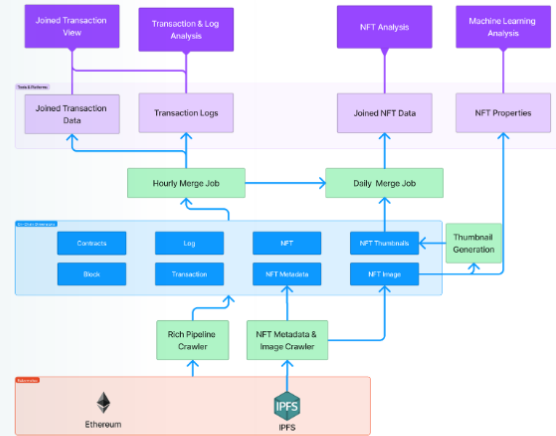
BAYC
#Token: 9.9K
#Transactions: 225K

Figure 7: Performance improvements of COMET compared to ALSTM cross different collections (5-step).

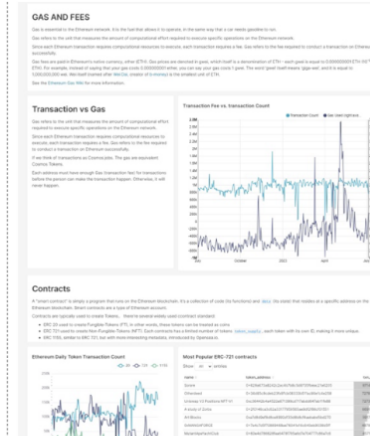
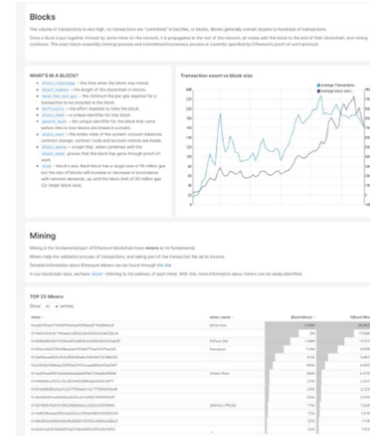
Practical Deployment

A

Establishment of data infrastructure



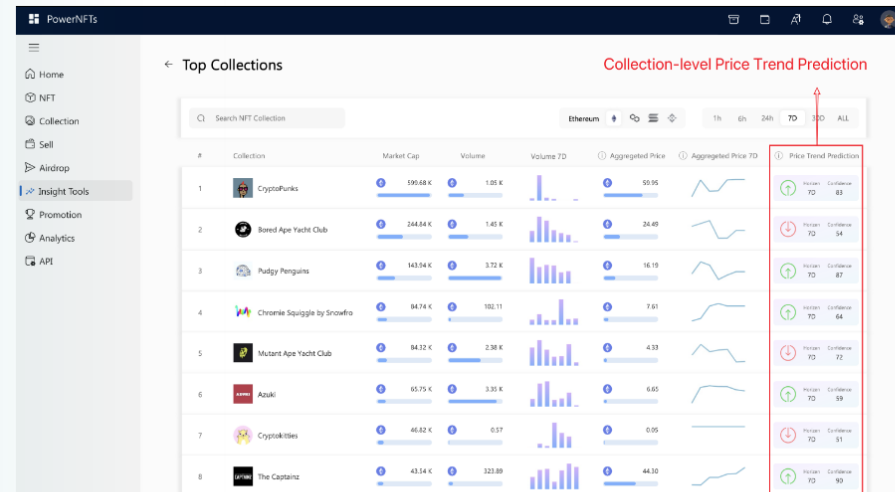
(a) Data collection pipeline



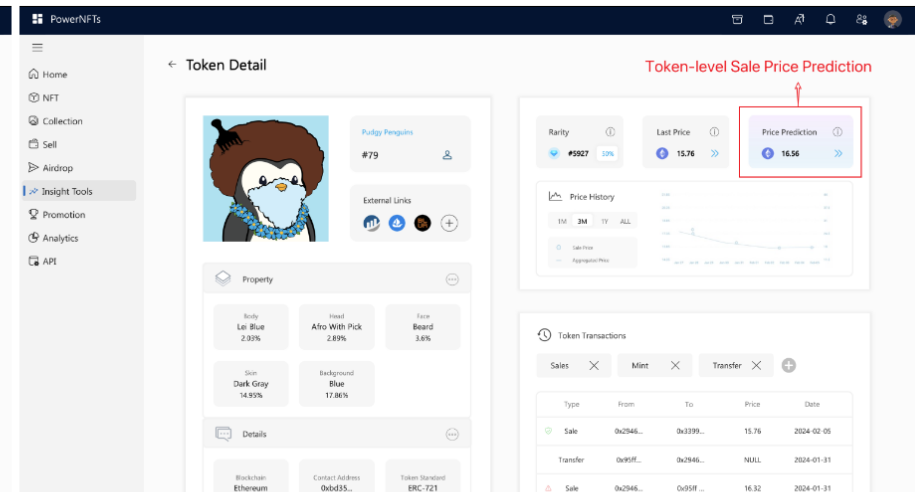
(b) Data visualization panel

B

Integration of practical products



(c) Top collections interface in PowerNFT



(d) Token details interface in PowerNFT

COMET: NFT Price Prediction with Wallet Profiling



Thanks

Tianfu Wang, Liwei Deng, Chao Wang, Jianxun Lian,
Yue Yan, Nicholas Jing Yuan, Qi Zhang, Hui Xiong



Paper



Code



Email

tianfuwang@mail.ustc.edu.cn

nicholas.yuan@microsoft.com

xionghui@ust.hk