



# NFT wash trading

Austen  
01.14.2023



# About Me

- 互联网金融行业 数据挖掘
- 定价、风控、用户转化
- web3 链上数据分析师
- 分析链上交易行为
- 后续计划
- 分享关于链上数据的一些数据分析、模型构建、训练思路(套利、安全、老鼠仓等)



**austen\_sfy**

@AustenSfy

Data scientist

Web3 beginner

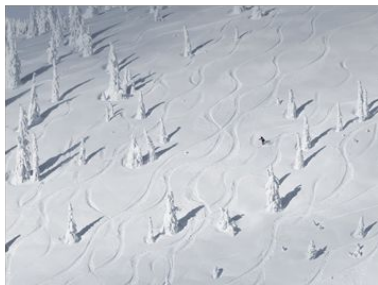
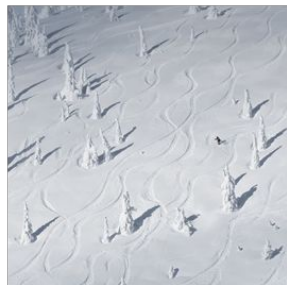
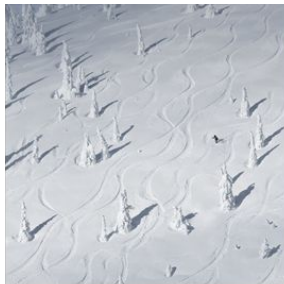
📅 Joined July 2022

# Contents



Footprint Analytics

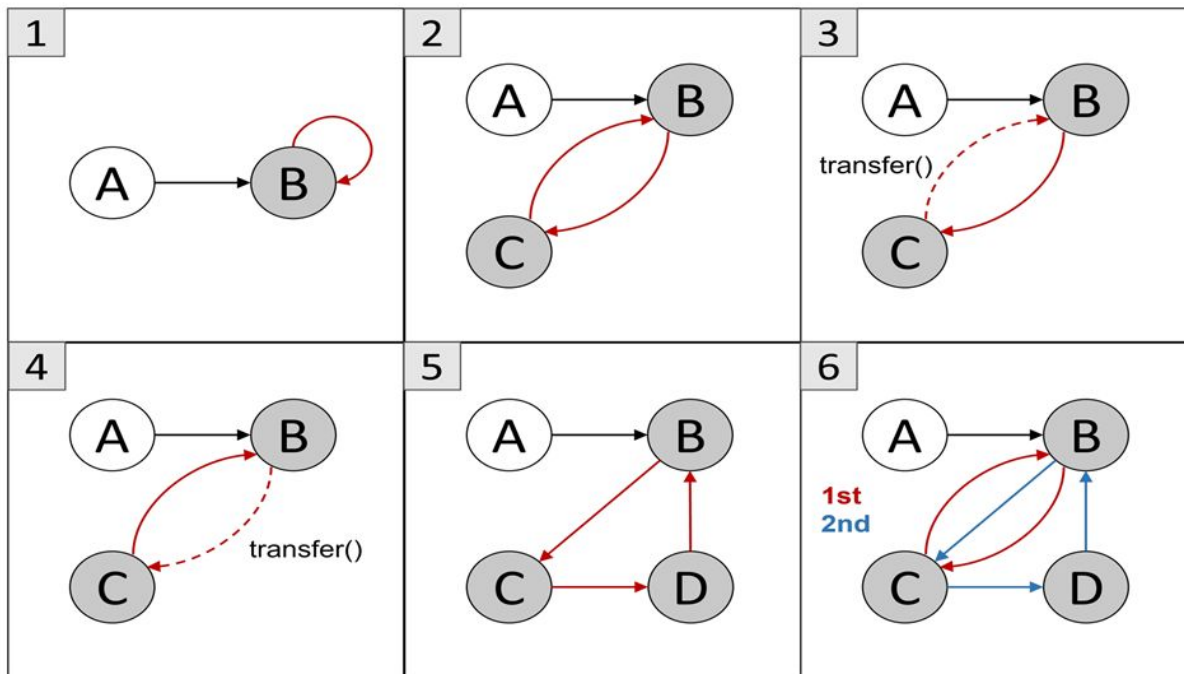
1. 什么是nft wash trading
2. 常见wash trading 类型
3. 实际事例(猴地)
4. 所需数据
5. 建模思路及特征构建
6. 模型训练
7. Footprint(规则)
8. 后续优化点



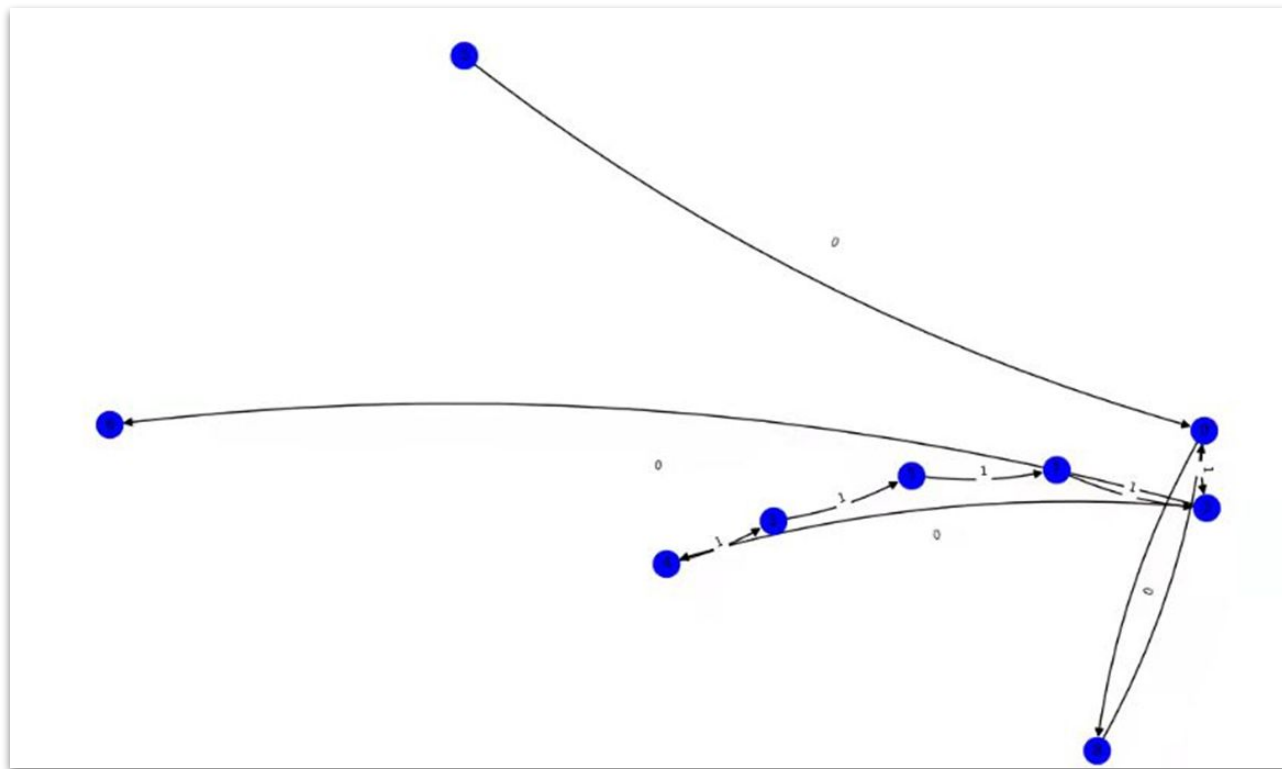
## 什么是nft wash trading

wash trading是一种误导买卖双方抬高价格的行为。买卖双方可以来回出售这件作品以抬高成本;作用:增加项目交易量、抬高价格

## 常见的nft wash trading操作方式



## 实际算法找出的事例



## 所需数据表

### Nft\_transactions

amount_currency_contract_address	amount_currency	amount
block_date	block_number	block_timestamp
buyer_address	chain	collection_contract_address
collection_slug	log_index	marketplace_contract_address
marketplace_slug	nft_token_id	number_of_nft_token_id
platform_fee_rate	platform_fees_amount	platform_fees_value
royalty_amount	royalty_rate	royalty_value
seller_address	trade_type	transaction_hash
value_currency	value	

### Nft\_transfers

amount_raw	block_date	block_number
block_timestamp	chain	from_address
internal_index	log_index	nft_token_id
to_address	token_address	transaction_hash
transfer_type		

# 特征构建

## 思路

以当前交易的nft\_id的线索, 追踪当前nft\_id后续10次的交易行为, 后续的因子构建均以10次交易行为进行(仅对猴地项目进行探索)

## 特征

- 1、去重后的买家钱包地址数量
- 2、去重后的卖家钱包地址数量
- 3、去重后的参与交易的钱包地址数量
- 4、本次交易的买家后续16次参与该nft\_id交易的次数(非转账)
- 5、本次交易的卖家后续16次参与该nft\_id交易的次数(非转账)



## 模型训练

将聚类算法中的离群点(样本)和孤立森林(样本)的异常值取交集, 得到异常交易数据

```
cluster_res_new = kmeans_outliers(feet_v_cp[feat_cols],5,False)
cluster_res_new = cluster_res_new.drop_duplicates()
cluster_res_new.head()
```

	sell_len	buy_len	buy_sell_len	sell_cnt	buy_cnt	Dist	OutLier
0	0.948278	0.385855	0.394407	-0.550331	-0.555741	8.158575	0
1	1.540634	1.602096	1.624030	-0.550331	-0.555741	19.291981	0
2	0.355922	0.385855	0.394407	-0.550331	-0.555741	6.448471	0
4	0.948278	0.993975	1.009219	-0.550331	-0.555741	11.771536	0
32	1.540634	0.385855	-0.220405	-0.550331	-0.555741	9.472172	0

```
from sklearn.ensemble import IsolationForest
model=IsolationForest(n_estimators=100, max_samples='auto', contamination=float(0.05))
model.fit(feet_v)
label2 = model.predict(feet_v)
nft_data_merge['label_iso'] = label2
nft_data_merge['label_iso'].value_counts()
```

```
1    236911
-1    11792
Name: label_iso, dtype: int64
```

# Footprint 早期规则

- 1、无论交易是否被高估，每个 NFT 都会比较其在 Opensea 上的历史平均价格。如果它大 [10] 倍，则标记为 1。
- 2、无论交易是否被高估，每个 NFT 都会比较其历史最低价。如果它大 [300] 倍，则标记为 1。
- 3、NFT一天内的总交易量是否大于[5]。如果是，则标记为 1。

# 后续优化点

- 1、尝试用footprint因子进行模型探索
- 2、复杂网络的环状结构进行探索
- 3、其他社区发现算法的尝试



# Welcome to join our community

---



Twitter: [https://twitter.com/Footprint\\_Data](https://twitter.com/Footprint_Data)



Discord: <https://discord.com/invite/3HYaR6USM7>



Data API: <https://www.footprint.network/data-api>



Github: <https://github.com/footprint-network>



Email: [sales@footprint.network](mailto:sales@footprint.network)



扫码添加Footprint小助手  
加入数据分析交流群~