

Alma Mater Studiorum - Università di Bologna  
Dipartimento Informatica - Scienza e Ingegneria - DISI  
Laurea Magistrale in Informatica  
Corso di Blockchain and Cryptocurrencies

# Analysis of the Complex Networks of Tweets Related to NFTs

Martina Daghia - 0001097932 - [martina.daghia@studio.unibo.it](mailto:martina.daghia@studio.unibo.it)

Martina Zauli - 0001097933 - [martina.zauli@studio.unibo.it](mailto:martina.zauli@studio.unibo.it)

Riccardo Spini - 0001084256 - [riccardo.spini@studio.unibo.it](mailto:riccardo.spini@studio.unibo.it)

*1 luglio 2024*

*Anno scolastico 2023-2024*

# Abstract

Nell'era dei social network, piattaforme come Facebook, Instagram e Twitter hanno ridefinito il panorama delle interazioni umane, influenzando sia il sociale che l'economia globale. Simultaneamente, il mondo degli NFT (Non-Fungible Token) ha conosciuto una rapida ascesa, offrendo nuove modalità per la creazione, il possesso e lo scambio di asset digitali unici e indivisibili. Questi sviluppi convergono nel nostro progetto, che si propone di esplorare la correlazione tra le attività e i pareri espressi sulla piattaforma di microblogging Twitter e le variazioni di prezzo degli NFT.

Il progetto si propone di esplorare le correlazioni tra le variazioni dei prezzi dei NFT (Non-Fungible Token) e l'analisi dei sentimenti dei tweet associati. Utilizzando dati storici di transazioni NFT e analisi dei sentimenti su Twitter, il progetto identifica i momenti di maggiore fluttuazione dei prezzi degli NFT e ricerca i tweet più influenti che potrebbero aver contribuito a tali variazioni. Attraverso tecniche di visualizzazione e analisi dei dati, vengono esaminati i picchi positivi e negativi nelle variazioni dei prezzi e viene valutata la correlazione con i sentimenti espressi sui social media. Questo studio mira a fornire una visione sulle dinamiche di mercato degli NFT e sull'importanza dell'opinione pubblica nell'ecosistema digitale emergente dei beni non fungibili.

# Indice

1	Descrizione del problema . . . . .	1
1.1	Obiettivo . . . . .	2
2	Dataset . . . . .	3
3	Analisi e risultati . . . . .	5
3.1	Analisi iniziale sui dati . . . . .	5
3.2	Analisi iniziale sull'andamento degli NFT . . . . .	7
3.3	Sentiment analysis . . . . .	9
3.4	Analisi variazione prezzo NFT . . . . .	11
3.5	Analisi correlazione tra i tweet e il prezzo degli NFT . . . . .	14
3.5.1	Analisi influenza tweet su NFT . . . . .	14
3.5.2	Analisi opinion leader . . . . .	20
3.5.3	Metodo silhouette . . . . .	21
3.5.4	Analisi influenza NFT sulla quantità di tweets . . . . .	25
4	Conclusioni . . . . .	31
4.1	Risultati ottenuti . . . . .	31
4.1.1	Risultati: se i tweet hanno influenzato gli NFT . . . . .	31
4.1.2	Risultati: se gli NFT hanno influenzato i tweet . . . . .	32
4.2	Limitazioni . . . . .	33
4.3	Lavori futuri . . . . .	34
5	Lavori correlati . . . . .	35

# Elenco delle figure

1	Datasets organizzati in due cartelle distinte per separare NFT e twitter. . . . .	3
2	Tweet con <b>hashtag pertinenti</b> agli NFT costituiscono più della metà della torta pari al <b>54,5%</b> (2.604.393 tweet) a confronto con quelli senza hashtag di interesse (che ammontano a 2.175.192 tweet). . . . .	6
3	Andamento dei tweet e degli NFT nel periodo da novembre 2020 a febbraio 2022. . .	8
4	Distribuzione della sentimental analysis (dal 01/12/2020 al 23/02/2022). . . . .	10
5	Distribuzione sentiment dei tweet, con percentuali e numeri. . . . .	11
6	Variazioni di prezzo degli NFT: picchi positivi e negativi. . . . .	12
7	Variazione per secondo in ordine decrescente per picchi positivi e negativi. . . . .	13
8	Distribuzione dei tweet con sentiment per ogni picco. . . . .	17
9	Distribuzioni relative ai picchi negativi dei tweet analizzati. . . . .	19
10	Grafici k-means relativi ai picchi positivi e negativi dei tweets analizzati. . . . .	22
11	Distribuzioni relative ai picchi negativi dei tweet analizzati. . . . .	23
12	Grafici per confrontare l'andamento della variazione del prezzo degli NFT e l'andamento della variazione del numero di tweet per ciascuno dei picchi considerati. . . .	26
13	Correlazioni di Pearson e p-value per ogni token NFT con variazione positiva . . . .	28

# Elenco delle tabelle

2	Risultati <b>iniziali</b> dei tweet antecedenti di tre giorni ad ogni picco positivo NFT. . .	15
3	Risultati dopo il <b>filtraggio</b> sulla media del punteggio totale. . . . .	16
4	Risultati <b>iniziali</b> dei tweets antecedenti di tre giorni al picco NFT negativo. . . .	19
5	Risultati dopo il <b>filtraggio</b> sulla media del punteggio totale del picco negativo. . . .	19
6	Tabella opinion leaders per ogni token. . . . .	23
7	Dati del tweet e descrizione dell'autore. . . . .	24
8	In questa tabella sono rappresentati i primi 5 picchi positivi dei token NFT e l'ultimo picco negativo (token 9032). Per ciascun token, sono indicati il numero di opinion leader di rilievo, il numero di tweets suddiviso per sentiment e la collezione dei tweets.	32

# 1 Descrizione del problema

Negli ultimi anni il mercato degli NFT è cresciuto in modo sbalorditivo. Questo nuovo token deriva dall'idea di un nuovo standard Ethereum che riesca a distinguere univocamente ogni token. Infatti, può essere associato a proprietà virtuali e digitali ed è in grado di identificare in modo univoco le proprietà a cui si riferisce.

Gli NFT sono una nuova tecnologia basata sulla blockchain che permette di collezionare oggetti in formato digitale attraverso il possesso di un token non fungibile. Più precisamente un non fungible token (NFT) è un token non fungibile memorizzato su una blockchain e che non è scambiabile con altre risorse digitali.

Parallelamente, i social media sono diventati una piattaforma centrale per la condivisione di idee, discussioni e tendenze culturali. Attraverso questi canali digitali, milioni di individui partecipano attivamente alla costruzione di narrazioni condivise e influenzano il destino di asset digitali come gli NFT.

Nel mondo dei social network, Twitter è uno dei più particolari e interessanti: si distingue per le sue funzioni differenti da quelle disponibili sulle altre piattaforme e, inoltre, mira maggiormente ad offrire un servizio di microblogging rispetto agli altri social, per cui risulta più informativo sulle vicende che accadono nel mondo.

Twitter viene creato da Jack Dorsey il 21 marzo 2006 a San Francisco. La piattaforma è spopolata fin da subito, diventando una delle più utilizzate tra i social network. Attualmente ha più di 300 milioni di utenti attivi, ciascuno dei quali scrive quotidianamente 65 milioni di tweet e processa più di 800000 richieste di ricerche ogni giorno [5].

Il nostro progetto mira a esplorare la complessa interazione tra gli NFT e il mondo dei social media, concentrandosi in particolare su Twitter. Siamo spinti dalla domanda intrigante: "Qual è la natura della relazione tra i prezzi degli NFT e il numero di tweet correlati?". L'obiettivo che ci poniamo è quello di capire la correlazione tra il prezzo di determinati NFT e la quantità di tweet correlati. In particolar modo, si vuole stabilire se l'aumento del prezzo degli NFT provoca un incremento del numero di tweet oppure se accade il contrario, ovvero se sono i tweet a influenzare

il prezzo. In secondo luogo, verrà svolta un'ulteriore analisi di rete per capire se l'effetto di rete è causato da specifici nodi centrali: per fare ciò, è stato necessario concentrarsi su una parziale visione della rete di Twitter, filtrando ed estrapolando solo gli account che generano i tweet.

## 1.1 Obiettivo

L'obiettivo di questo studio è investigare la relazione tra i tweet e gli NFT: nello specifico, vogliamo comprendere se i tweet influenzano l'andamento dei prezzi degli NFT (Non-Fungible Token), oppure se sono gli NFT ad incidere in qualche misura sui tweet degli utenti.

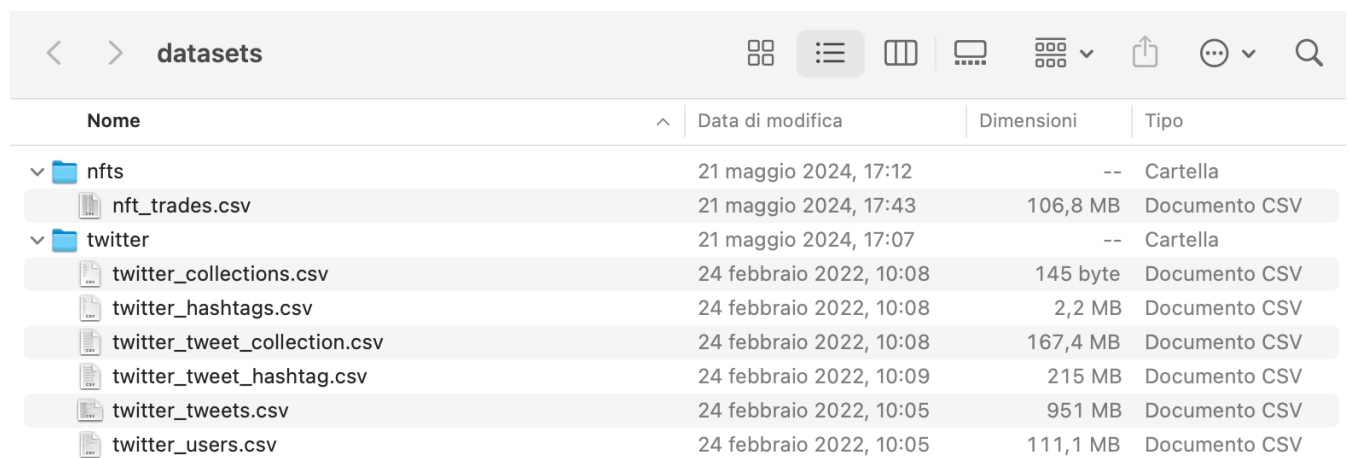
Per raggiungere questo obiettivo, il progetto si propone di:

- Analizzare un ampio set di dati contenente informazioni dettagliate sui tweet e le transazioni degli NFT;
- Identificare specifici hashtag, parole chiave o temi associati ai tweet che potrebbero influenzare il mercato degli NFT ed esaminarne il contenuto (sentiment analysis, topics, ecc...);
- Identificare i tweet più influenti e gli strong opinion leader;
- Individuare i picchi delle variazioni di prezzo degli NFT;
- Misurare la correlazione temporale tra l'emissione di tweets virali e le variazioni di prezzo degli NFT nelle ore e nei giorni successivi ai tweet (e viceversa);
- Valutare, infine, se esiste una tendenza in cui i tweet influenzano i movimenti di prezzo degli NFT o se sono le fluttuazioni di prezzo a generare discussioni tra gli utenti su Twitter.

## 2 Dataset

I datasets sono contenuti all'interno della cartella "datasets" collocata al di fuori della cartella del progetto (questo perché contiene file di dimensioni troppo grandi per poter essere caricata su github). In seguito, i file CSV sono stati salvati in ulteriori sottocartelle nominate "nfts" e "twitter".

Durate le analisi svolte che vedremo in seguito, è stato necessario salvare i dataframes in formato CSV creati durante l'implementazione del progetto: questi nuovi file sono stati salvati nella cartella "dataframes", anch'essa creata e mantenuta al di fuori della cartella del progetto (proprio come la cartella "datasets").



Nome	Data di modifica	Dimensioni	Tipo
▼ nfts	21 maggio 2024, 17:12	--	Cartella
nft_trades.csv	21 maggio 2024, 17:43	106,8 MB	Documento CSV
▼ twitter	21 maggio 2024, 17:07	--	Cartella
twitter_collections.csv	24 febbraio 2022, 10:08	145 byte	Documento CSV
twitter_hashtags.csv	24 febbraio 2022, 10:08	2,2 MB	Documento CSV
twitter_tweet_collection.csv	24 febbraio 2022, 10:08	167,4 MB	Documento CSV
twitter_tweet_hashtag.csv	24 febbraio 2022, 10:09	215 MB	Documento CSV
twitter_tweets.csv	24 febbraio 2022, 10:05	951 MB	Documento CSV
twitter_users.csv	24 febbraio 2022, 10:05	111,1 MB	Documento CSV

Figura 1: Datasets organizzati in due cartelle distinte per separare NFT e twitter.

Abbiamo effettuato delle operazioni di cleaning per alleggerire il dataset degli NFT, eliminando gli attributi "contract\_address" e "currency\_address" che non erano fondamentali per la nostra analisi.

I datasets usati sono i seguenti con le relative variabili:

### 1. Tweets

File: `'twitter_tweet.csv'`

Contiene: `tweet_id`, `created_at`, `author_id`, `content`, `retweet_count`, `reply_count`, `like_count`, `quote_count`



## 2. Twitter Users

File: '*twitter\_users.csv*'

Contiene: `user_id`, `name`, `username`, `description`

## 3. NFT Trades

File: '*nft\_trades.csv*'

Contiene: `date`, `transaction_hash`, `num_items`, `seller`, `buyer`, `token_id`, `avg_price`, `currency`, `usd_amount`, `platform`

## 4. Twitter-Tweet Hashtag Relationship

File: '*twitter\_tweet\_hashtag.csv*'

Contiene: `tweet_id` e `hashtag`

## 5. Twitter-Tweet Collection Relationship

File: '*twitter\_tweet\_collection.csv*'

Contiene: `tweet_id` e `collection`

## 6. Twitter-Tweet Collection

File: '*twitter\_collections.csv*'

Contiene: `collection`

## 7. Twitter-Tweet Collection

File: '*twitter\_hashtags.csv*'

Contiene: `hashtags`

Inizialmente, la mole di dati era composta da **507678** entries inerenti agli NFT (estrapolati dal file *nft\_trades.csv*) e da **4792264** entries riguardanti i tweets (estrapolati dal file *twitter\_tweets.csv*).

## 3 Analisi e risultati

Per garantire maggiore chiarezza, abbiamo suddiviso tutti i passaggi logici effettuati in sottosezioni, descrivendo nel dettaglio ciascun passo. È importante sottolineare che, in seguito a ciascuna importazione dei dataset e all'inizio di ogni analisi, sono state fatte ulteriori operazioni di pulizia per eliminare i valori nulli e quelli duplicati.

### 3.1 Analisi iniziale sui dati

Per prima cosa, abbiamo deciso di assegnare un peso ad alcune variabili del file "*twitter\_tweet.csv*" per poter individuare i tweet più virali.

Abbiamo stabilito che i **like** del tweet avessero minor importanza rispetto agli altri fattori, perciò abbiamo attribuito ad essi un peso pari a 1. Alle **reply** abbiamo assegnato un peso pari a 2. Infine, ci siamo concentrati maggiormente sulla valutazione dei **retweet** e dei **quote**, assegnando loro un peso di 3 e 4 rispettivamente.

Abbiamo attribuito un'importanza maggiore ai quote e ai retweet rispetto ai like e le reply perché riteniamo che, quando un utente decide di ri-condividere un tweet ripostandolo sul proprio profilo, pone molta più enfasi sull'argomento e ha necessità di condividerlo con i propri seguaci.

In particolare, la differenza tra i **retweet** e una **citazione** su Twitter risiede proprio nel contenuto [1]:

- Un **retweet** è quando un utente decide di condividere un tweet di qualcun altro direttamente sul proprio feed senza aggiungere alcun commento personale: in questo modo il contenuto e l'autore del tweet originale sono chiaramente visibili.
- Una **citazione**, o "**quote tweet**", è quando un utente decide di condividere un tweet di qualcun altro nel proprio feed aggiungendo un proprio commento o messaggio personale, includendo il contenuto del tweet originale. Questo permette all'utente di condividere il tweet originale e di aggiungere contesto, opinioni, domande o qualsiasi altro commento rilevante.

Abbiamo individuato la data del **primo tweet** (2017-01-06 21:42:43) e dell'**ultimo** (2022-02-23 23:57:21) per poi mostrare graficamente (3a) quali siano stati i mesi e gli anni con i tweet più influenti.

Abbiamo confrontato il grafico con le vendite degli NFT 3b (spiegazione più approfondita nella Sezione 3.2) e abbiamo notato che entrambi i grafici presentavano i picchi più alti circa nello stesso periodo. Pertanto, abbiamo deciso di **filtrare i tweet dal 1/12/2020 fino all'ultimo tweet** (Febbraio 2022).

Successivamente, dopo aver importato il dataset "*twitter\_tweet\_hashtag.csv*", abbiamo redatto una lista di hashtag relativi agli NFT e li abbiamo visualizzati graficamente. Abbiamo **filtrato i primi 25 hashtag** e salvato i tweet contenenti questi hashtag. Abbiamo poi proceduto con la sentiment analysis su questi tweet.

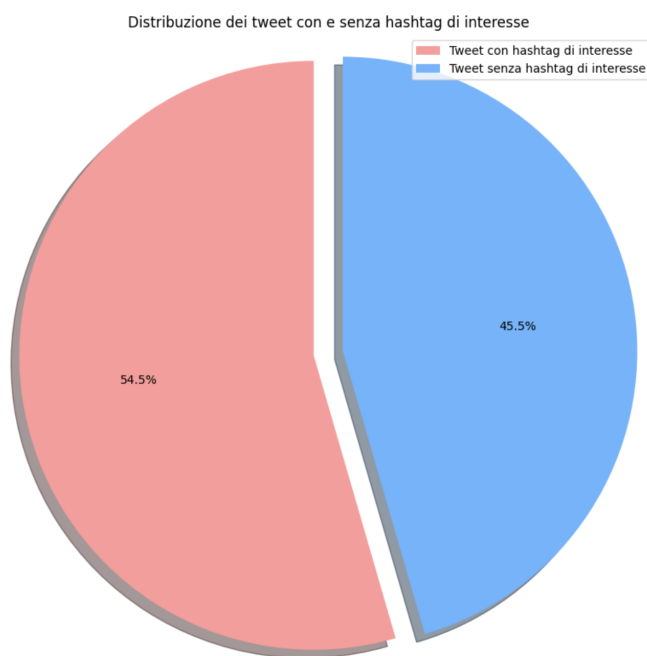
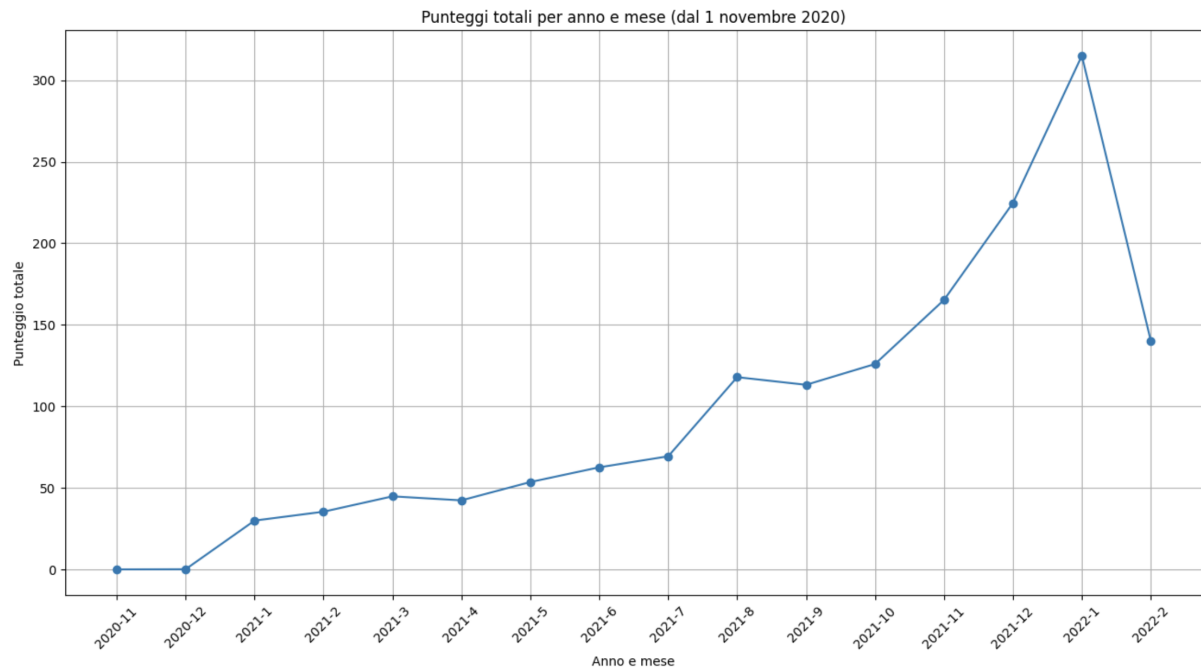


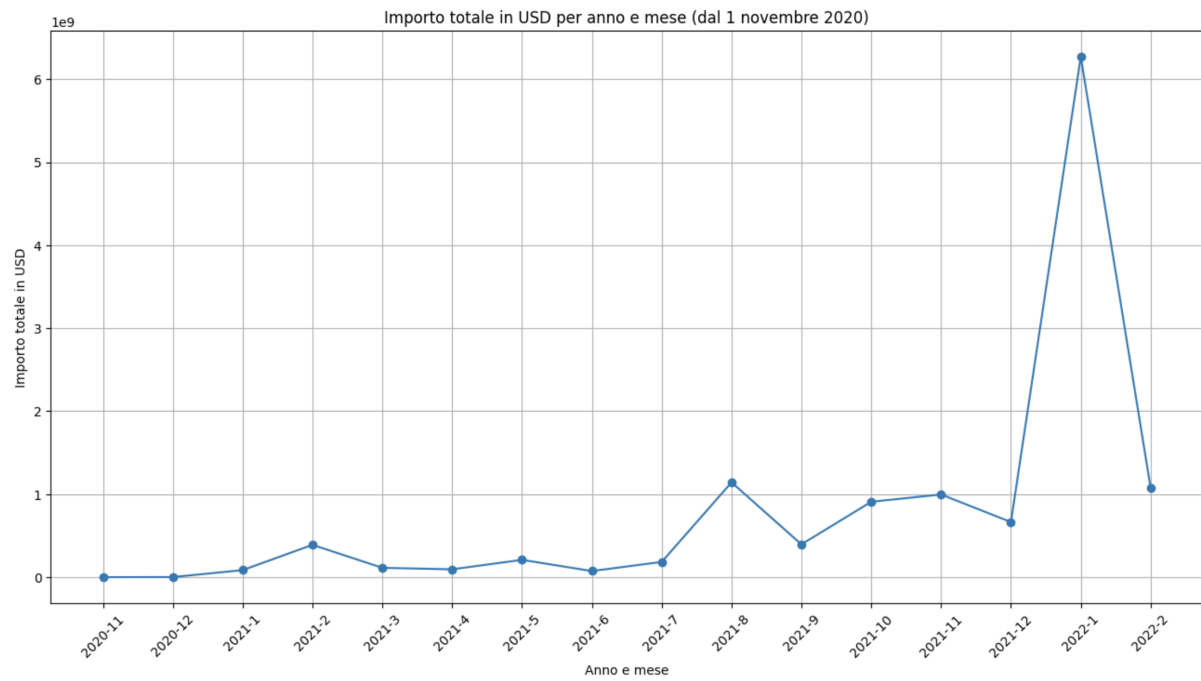
Figura 2: Tweet con **hashtag pertinenti** agli NFT costituiscono più della metà della torta pari al **54,5%** (2.604.393 tweet) a confronto con quelli senza hashtag di interesse (che ammontano a 2.175.192 tweet).

## 3.2 Analisi iniziale sull'andamento degli NFT

Per identificare i **periodi di interesse** sull'andamento dei prezzi degli NFT, abbiamo creato un grafico che copre l'intero intervallo di date del dataset (vedi Figura 3b). Successivamente, abbiamo confrontato questo grafico con quello relativo al numero di tweet (Figura 3a), al fine di **individuare i picchi** coincidenti da analizzare più approfonditamente. Questa fase ci ha consentito di filtrare ulteriormente il dataset, un passaggio cruciale considerando le sue dimensioni iniziali: questo è stato fatto per porre il focus su un periodo che evidenziasse contemporaneamente un aumento significativo del numero di tweet e dei prezzi degli NFT.



(a) Andamento dei tweet per **punteggio totale** nel periodo da novembre 2020 a febbraio 2022.



(b) Andamento del **prezzo totale** degli NFT nel periodo da novembre 2020 a febbraio 2022.

Figura 3: Andamento dei tweet e degli NFT nel periodo da novembre 2020 a febbraio 2022.

### 3.3 Sentiment analysis

La sentiment analysis è uno strumento del Natural Language Processing (NLP) che serve a determinare l'emotività e il sentimento espresso in un testo.

Noi abbiamo sfruttato `TextBlob`, una libreria Python per l'elaborazione del linguaggio naturale che consente di effettuare operazioni sofisticate sui dati testuali. `TextBlob` è in grado di **analizzare un testo e definirne il sentimento** in base all'orientamento semantico e all'intensità di ogni parola in una frase. Questa libreria assegna punteggi individuali a tutte le parole contenute nel testo e calcola un sentimento finale [4].

Grazie a questa libreria, siamo riusciti ad estrapolare le emozioni degli utenti nei loro tweet.

Analizzare e capire il sentiment degli utenti ci sarà utile per **capire le loro emozioni e/o opinioni negative, positive e neutrali nei confronti degli NFT** e capire, in seguito, se questo ha influenzato l'andamento del prezzo.

#### Implementazione

Trattando dati non strutturati, come i testi dei tweets, abbiamo ritenuto necessario pulire i dati della colonna "content" del dataset per ottenere un testo uniforme e pronto per analisi approfondite. Questo processo ha incluso la rimozione di determinati caratteri speciali, la tokenizzazione e l'eliminazione delle stopwords che non aggiungono valore al sentiment.

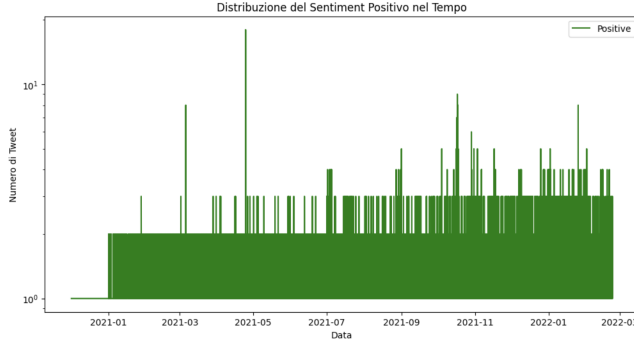
Dopo questa fase di pre-elaborazione, abbiamo calcolato la **polarità**, un valore compreso tra -1 e 1 che indica se il sentimento del tweet è **positivo**, **negativo** o **neutro**. Questo passaggio è stato essenziale per decifrare il sentiment degli utenti e ottenere un'analisi accurata delle opinioni espresse nei tweets.

Per analizzare i risultati, è importante sapere i criteri di classificazione dei tweets:

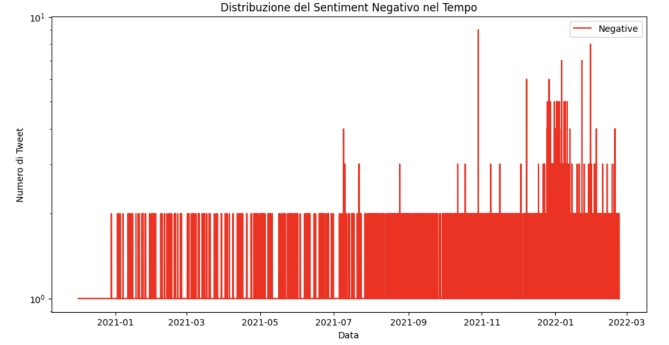
- **Polarità pari a 0:** il sentiment è neutro;
- **Polarità positiva:** il sentiment è positivo, con valori più vicini a 1 che indicano un sentiment molto positivo;

- **Polarità negativa:** il sentiment è negativo, con valori più vicini a -1 che indicano un sentiment molto negativo.

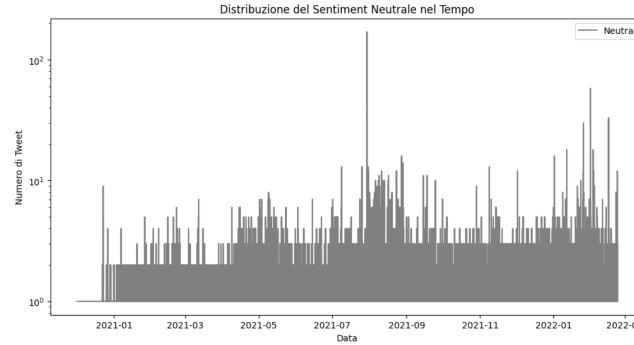
Abbiamo, poi, esportato il dataframe su un file in formato CSV, in modo tale da poter continuare l'analisi come spiegato alla sezione 3.5. I risultati della sentiment analysis è descritta come segue in figura 4, la quale mostra la quantità di tweet positivi (verde), negativi (rosso) e neutri (grigio):



(a) Distribuzione sentiment **positivo**.



(b) Distribuzione sentiment **negativo**.



(c) Distribuzione sentiment **neutrale**.

Figura 4: Distribuzione della sentimental analysis (dal 01/12/2020 al 23/02/2022).

Il rapporto complessivo tra i tweet che parlano di NFT con sentiment positivo, negativo e neutro è rappresentato dettagliatamente nella figura 5.

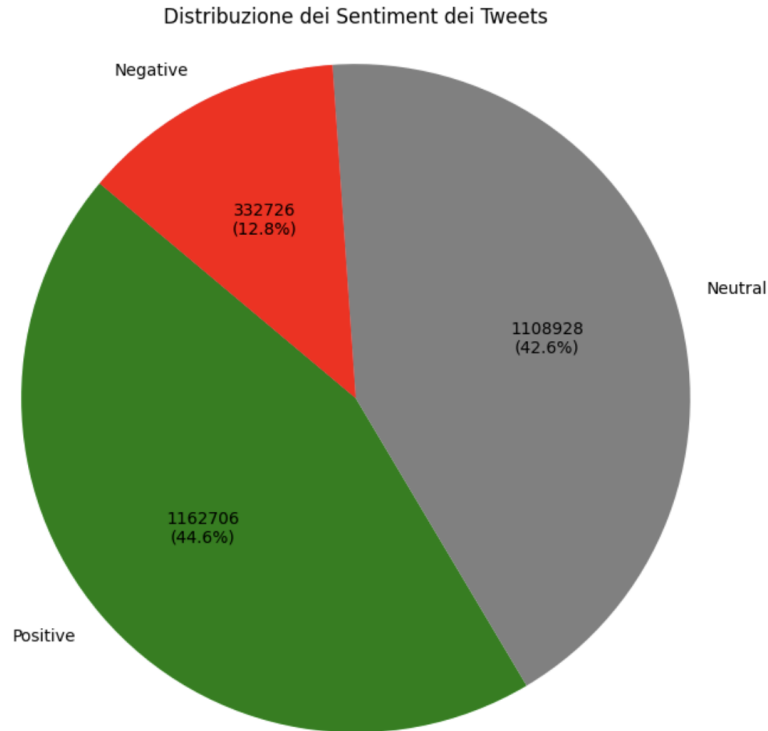


Figura 5: Distribuzione sentiment dei tweet, con percentuali e numeri.

### 3.4 Analisi variazione prezzo NFT

Per analizzare la variazione dei prezzi degli NFT, abbiamo usato il file dataset "*nft\_trades.csv*", prendendo in esame solo le transazioni avvenute dal 1 dicembre 2020 (inizio del periodo interessante di variazioni, come rilevato dall'analisi di cui parlavamo nella sezione 3.2).

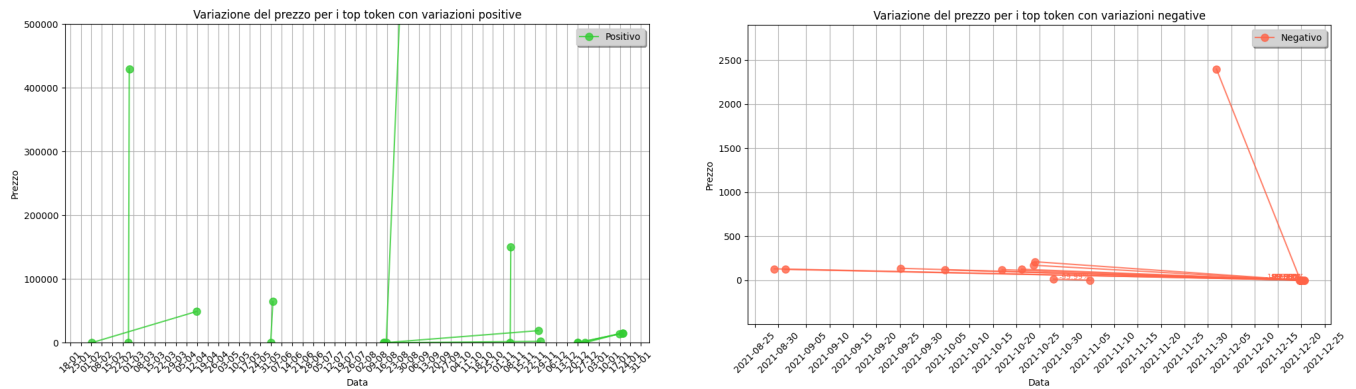
Innanzitutto, abbiamo raggruppato le transazioni per ogni token, identificando le variazioni di prezzo **positive** e **negative** consecutive di ciascun token presente nel dataset. Per ogni token che ha mostrato una variazione di prezzo significativa (escludendo quelli in cui la somma del prezzo della prima e dell'ultima transazione risultava uguale a 0), abbiamo calcolato l'inizio e la fine della variazione di prezzo in una certa direzione (positiva o negativa). Successivamente, abbiamo



determinato la percentuale di variazione del prezzo, prendendo in considerazione solo quei token che superavano una certa soglia di variazione percentuale.

Abbiamo quindi rapportato questa percentuale alla durata della variazione, identificando i periodi di inizio e fine della stessa. Questo ci ha permesso di calcolare la variazione percentuale del prezzo in relazione al tempo. Graficamente, ci riferiamo ai "picchi" come quei token che mostrano la variazione percentuale per unità di tempo più elevata, come illustrato in figura 10.

Questo metodo ci ha consentito di escludere i normali andamenti di mercato degli NFT, non rilevanti per il nostro studio. Identificando le variazioni più significative, siamo riusciti a individuare i token con le variazioni di prezzo più rapide e rilevanti, sia in positivo che in negativo.



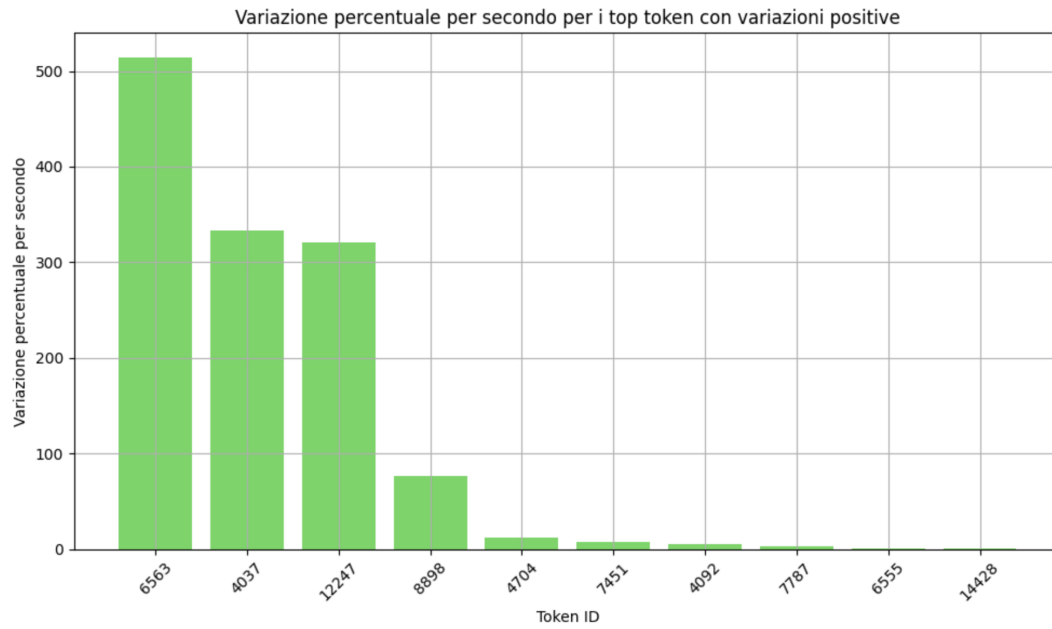
(a) **Picchi positivi** delle variazioni di prezzo degli NFT.

(b) **Picchi negativi** delle variazioni di prezzo degli NFT.

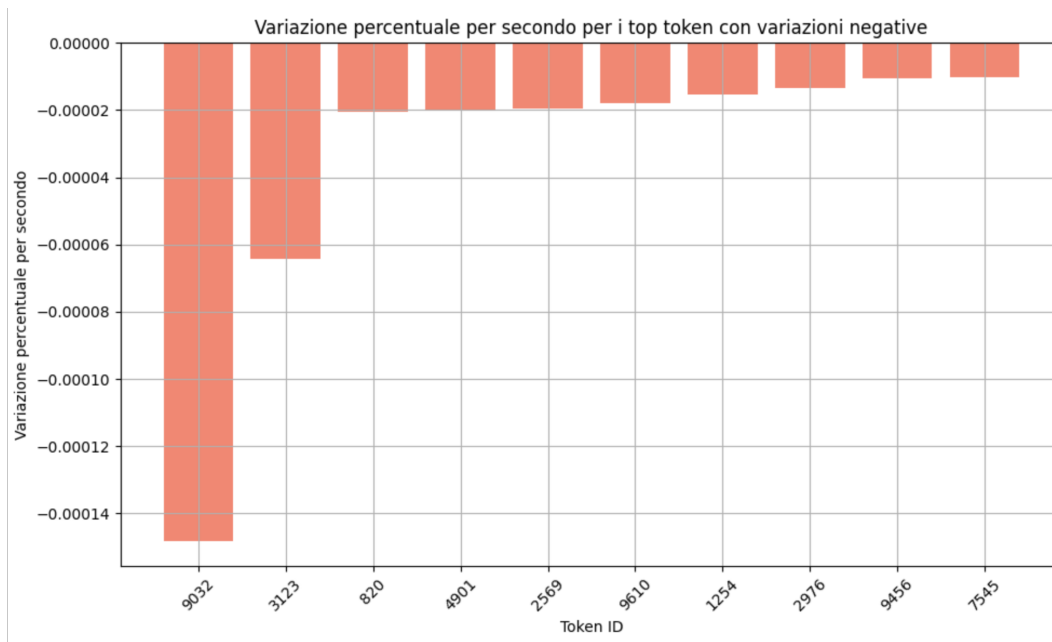
Figura 6: Variazioni di prezzo degli NFT: picchi positivi e negativi.

Infine, abbiamo deciso di selezionare i **5 picchi più alti** in termini di variazione al secondo per gli NFT con variazioni positive, e l'**unico picco negativo**, poiché rappresentava la variazione al secondo più alta e il minor intervallo di tempo tra la data di inizio e di fine.

Osservando il grafico 7b, abbiamo deciso di esaminare il picco con la maggiore variazione percentuale per secondo, ovvero il picco *9032*. Abbiamo scelto di non considerare altri picchi negativi poiché si estendevano su periodi di inizio e fine variazione più lunghi.



(a) Token positivi con la **maggior percentuale di variazione per sec.**



(b) Token negativi con la **maggior percentuale di variazione per sec.**

Figura 7: Variazione per secondo in ordine decrescente per picchi positivi e negativi.

Nelle tabelle di seguito mostriamo le **caratteristiche principali dei picchi** che abbiamo individuato e selezionato per procedere con le analisi successive.

### Picchi NFT risultanti

Data inizio	Data fine	Token ID	Prezzo inizio	Prezzo fine	Variazione per secondo
2021-08-15 07:05:16	2021-08-30 16:33:32	<i>6563</i>	0.13	888888.0	514.07
2021-11-05 09:46:35	2021-11-05 18:42:28	<i>4037</i>	1.4	150000.0	333.23
2021-02-25 14:23:29	2021-02-26 02:48:36	<i>12247</i>	3.0	430000.0	320.60
2021-05-30 18:06:57	2021-06-01 06:39:14	<i>8898</i>	0.64	64842.0	77.02
2021-12-20 02:29:35	2022-01-18 23:35:20	<i>4704</i>	0.05	15000.0	11.62
2021-10-27 23:46:06	2021-11-04 19:14:53	<i>9032</i>	176183	0.001	-0.000148

## 3.5 Analisi correlazione tra i tweet e il prezzo degli NFT

Nell'ambito di questa analisi, è stato eseguito uno studio sull'influenza dei tweet sui picchi di interesse per i token NFT. Sono stati identificati e analizzati sia i picchi positivi che quelli negativi, confrontando l'impatto dei tweet positivi e negativi sui movimenti del mercato NFT.

### 3.5.1 Analisi influenza tweet su NFT

Per riuscire a soddisfare il nostro primo obiettivo, ovvero quello di capire se ci sono stati tweet che hanno influenzato la variazione degli NFT, abbiamo analizzato i picchi positivi e negativi degli NFT (come visto nella sezione precedente 3.4), i tweet con il relativo sentiment e la collezione dei tweet.

Originariamente, abbiamo analisi per primi i picchi positivi e successivamente quello negativo. Vediamo ora più nel dettaglio quali sono stati i passaggi effettuati.

### Analisi picchi positivi

Inizialmente, abbiamo individuato i tweet che sono stati creati i tre giorni antecedenti alla data di inizio della variazione dei picchi positivi degli NFT; in seguito, abbiamo calcolato per ogni picco NFT quanti sono stati i tweet effettuati e quanti di essi possiedono un sentiment positivo, negativo e neutrale.

Per prendere i tweet più influenti, abbiamo calcolato per ogni picco la media del punteggio totale (3.1) e preso i tweet con **punteggio totale maggiore stretto rispetto alla media**. Successivamente, abbiamo ricalcolato il numero di tweet totale separando anche quanti di questi hanno un sentiment positivo, negativo e neutrale.

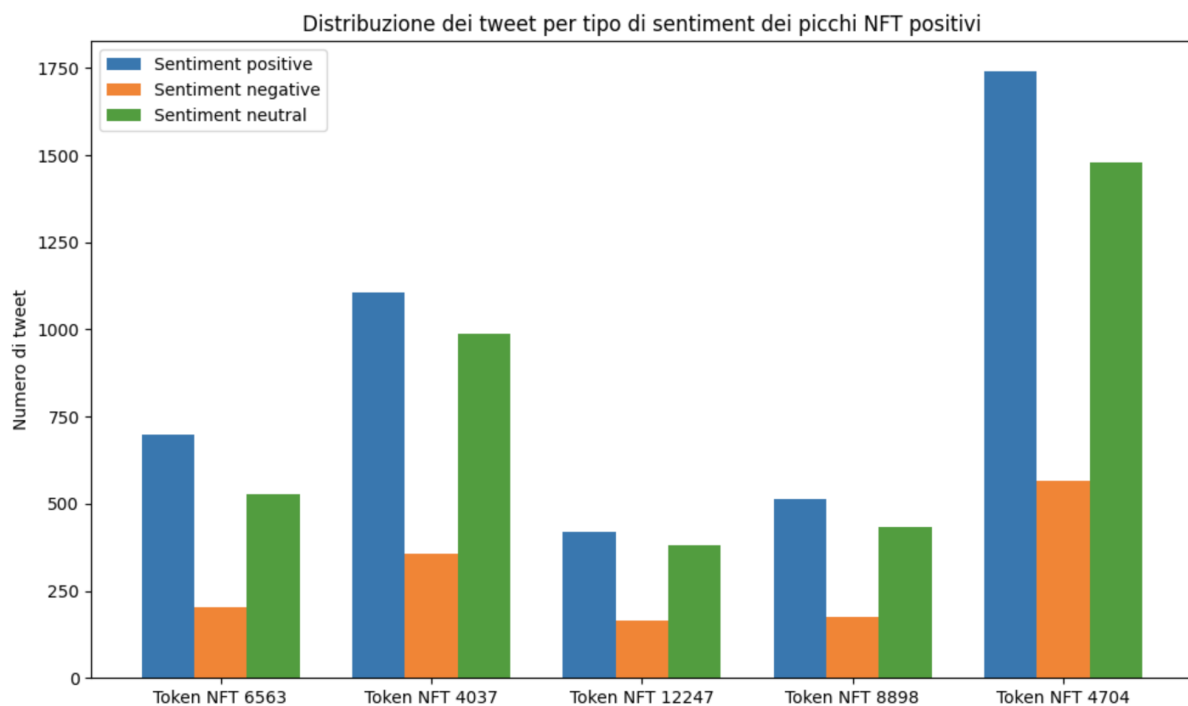
NFT Token	N° totale di tweet	Positivi	Neutrali	Negativi
6563	14026	6215	5880	1931
4037	23987	11717	9438	2832
12247	7083	3139	2617	1327
8898	8338	3905	3299	1134
4704	35917	15709	16275	3933

Tabella 2: Risultati **iniziali** dei tweet antecedenti di tre giorni ad ogni picco positivo NFT.

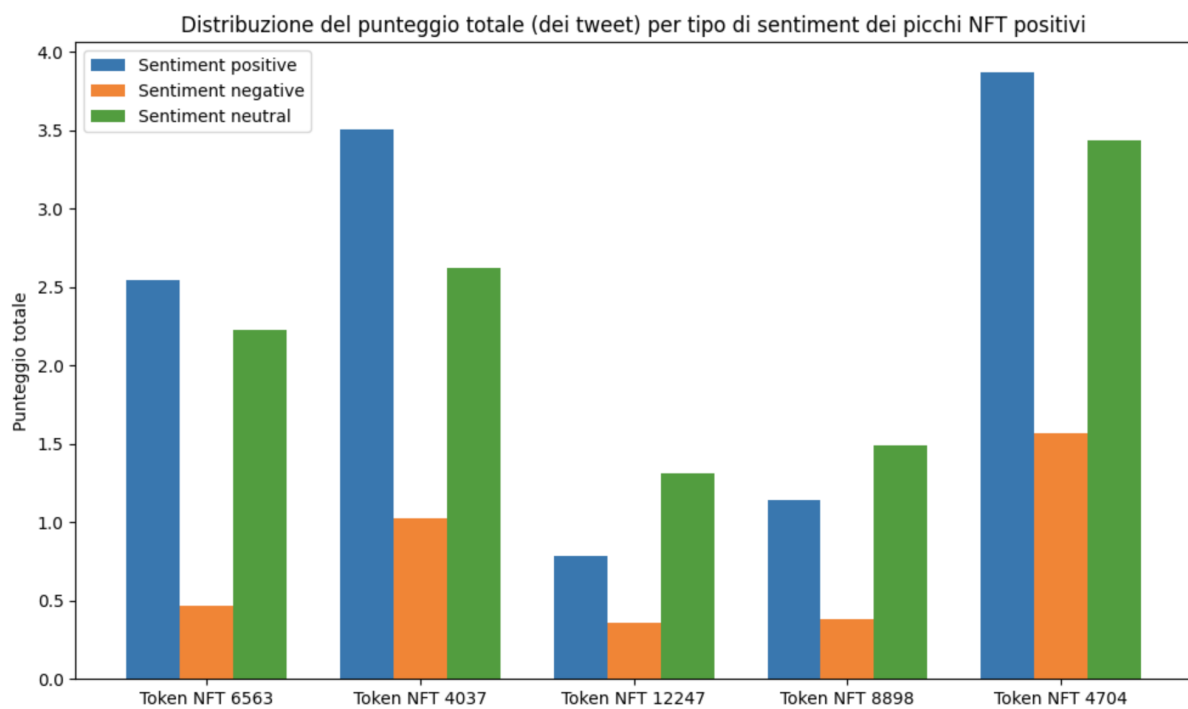
<b>NFT Token</b>	<b>Media punteg. tot.</b>	<b>N° tweet sopra media</b>	<b>Positivi</b>	<b>Neutrali</b>	<b>Negativi</b>
<i>6563</i>	0.0004	1430	697	529	204
<i>4037</i>	0.0003	2451	1106	987	358
<i>12247</i>	0.0004	967	421	382	164
<i>8898</i>	0.0004	1121	512	433	176
<i>4704</i>	0.0003	3786	1740	1480	566

Tabella 3: Risultati dopo il **filtraggio** sulla media del punteggio totale.

Abbiamo creato grafici a barre per ogni picco, illustrando la distribuzione dei tweet in base al tipo di sentiment (positivo, negativo e neutrale) e il punteggio totale dei tweet per ciascun tipo di sentiment.



(a) Distribuzione dei tweet per tipo di sentiment (NFT positivi).



(b) Distribuzione del punteggio totale per tipo di sentiment (NFT positivi).

Figura 8: Distribuzione dei tweet con sentiment per ogni picco.

Dai grafici 8a e 8b notiamo che il picco del token NFT *4704* possiede un numero di tweet elevato con sentiment positivo, seguito dai token *6563* e *4037*. Dopo aver filtrato i tweets virali basandoci sul loro punteggio totale, è possibile notare che ci sono molti tweets influenti con sentiment positivo, seguito dal sentiment neutrale.

Dopo aver esaminato i tweets filtrati sopra la media del punteggio totale, li abbiamo uniti con il dataset delle collezioni per analizzare le collezioni dei tweets.

Di conseguenza, abbiamo presupposto che quel picco di NFT sia collegato con la collezione più alta dei tweets. Abbiamo, quindi, contato i tweets per ogni collezione e visualizzato i risultati in un grafico a barre notando che:

- Per il picco positivo con l'ID del token **6563** la collezione che possiede il numero più alto di tweet è la **doodles** con 810 tweets;
- Per il picco positivo con l'ID del token **4037** la collezione che possiede il numero più alto di tweet è la **doodles** con 1267 tweets;
- Per il picco positivo con l'ID del token **12247** la collezione che possiede il numero più alto di tweet è la **doodles** con 843 tweets;
- Picco positivo con il token ID **8898** la collezione che possiede il numero più alto di tweet è la **doodles** con 770 tweets;
- Per il picco positivo con l'ID del token **4704** la collezione che possiede il numero più alto di tweet è la **doodles** con 1756 tweet, ma anche la collezione **boredapeyachtclub** è alta con 1285 tweets.

### Analisi picco negativo

Per gestire il picco negativo del token NFT abbiamo proceduto come per i picchi positivi eseguendo gli stessi passaggi.

NFT Token	N° totale di tweets	Positivi	Neutrali	Negativi
9032	24705	11272	10869	2564

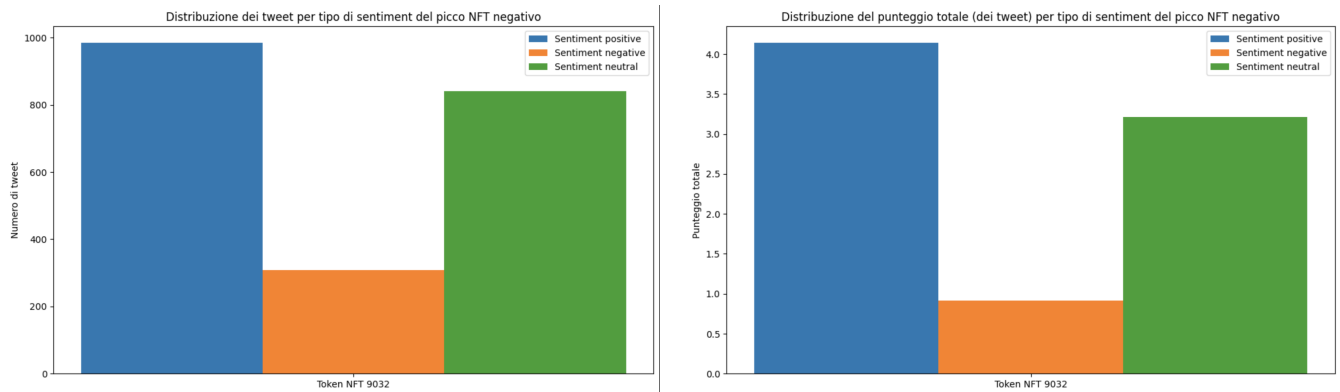
Tabella 4: Risultati **iniziali** dei tweets antecedenti di tre giorni al picco NFT negativo.

NFT Token	Media punteg. tot.	N° tweets sopra media	Positivi	Neutrali	Negativi
9032	0.0004	2133	984	840	309

Tabella 5: Risultati dopo il **filtraggio** sulla media del punteggio totale del picco negativo.

In quest'ultima tabella, notiamo che i tweets negativi antecedenti di tre giorni all'inizio della data di inizio variazione del token sono comunque bassi rispetto a quelli positivi e neutrali.

Dai grafici 9a e 9b, abbiamo notato che questo picco ha una **bassa influenza negativa** e abbiamo presupposto che potrebbe essere stato influenzato da **altri fattori esterni** e non dai tweets.



(a) Distribuzione dei tweet per tipo di sentiment del picco negativo.

(b) Distribuzione del punteggio totale per tipo di sentiment del picco negativo.

Figura 9: Distribuzioni relative ai picchi negativi dei tweet analizzati.



### 3.5.2 Analisi opinion leader

Per procedere con l'analisi **opinion leader**, abbiamo tenuto in considerazione dei tweets filtrati nella sezione precedente ed abbiamo esportato per ogni picco i tweets che ci interessavano.

Abbiamo utilizzato **k-means**, un algoritmo di clustering utilizzato nell'apprendimento automatico non supervisionato per suddividere un determinato set di dati in  $k$  cluster in base alla somiglianza delle loro caratteristiche. L'algoritmo funziona assegnando iterativamente i punti dati al centroide del cluster più vicino e aggiornando i centroidi in base alla media dei punti dati assegnati al cluster. Il processo continua fino a quando i centroidi non si spostano più in modo significativo o viene raggiunto un numero massimo di iterazioni. La spiegazione sull'importanza che abbiamo dato a like, reply, retweets e quote è riportata nella sezione 3.1.

Una volta individuati gli strong opinion leader grazie alla rappresentazione grafica del k-means, abbiamo voluto visualizzare l'elenco di questi utenti e stamparne i relativi tweets andati virali: in questo modo, abbiamo potuto indagare sull'influenza che questi tweets hanno avuto sulla variazione di prezzo degli NFTs.

Per determinare chi sono gli opinion leaders, abbiamo utilizzato parametri che misurano il grado di interazione con un tweet: like, retweet, risposte e citazioni. Abbiamo calcolato questi parametri per ogni singolo tweet e poi abbiamo calcolato quanti like, retweet, risposte e citazioni ha ricevuto ogni account durante i **periodi critici**. I periodi critici che abbiamo considerato vanno da 3 giorni prima dell'inizio della variazione dell'NFT, fino alla sua data di fine variazione.

Il nostro lavoro per la ricerca dell'opinion leader è stato svolto in una serie di dataframes che descrivono l'insieme di tweets postati 3 giorni prima dell'inizio della variazione degli NFT in questione, filtrandoli in base ad una **soglia sul punteggio** (abbiamo preso in considerazione tutti i tweets con punteggio superiore alla media dei punteggi totali per ogni picco). In questo modo i tweets con cui abbiamo lavorato erano già filtrati per una certa importanza.

### 3.5.3 Metodo silhouette

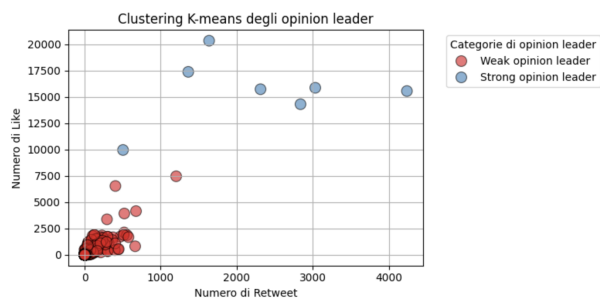
L'algoritmo k-means richiede di specificare in anticipo il numero di cluster ( $k$ ), che può essere determinato attraverso tecniche come l'analisi della silhouette.

In questo lavoro, consideriamo il punteggio Silhouette, una metrica che misura quanto ogni punto dati si adatti al cluster assegnato rispetto agli altri cluster. Quest'ultima è una misura di quanto bene i punti di un cluster sono separati rispetto agli altri cluster, variando da -1 a 1.

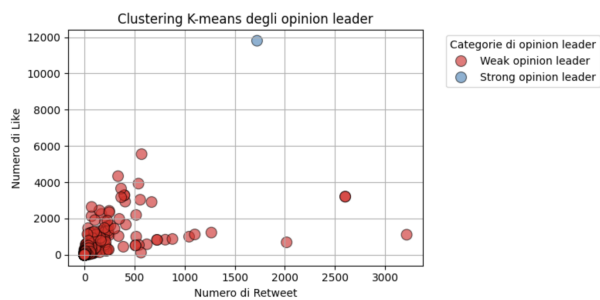
In tutti i picchi (positivi e negativi), abbiamo suddiviso i dati in due classi, corrispondenti a due cluster: **Strong Opinion Leaders** e **Weak Opinion Leaders**.

I grafici sottostanti illustrano i cluster che identificano un opinion leader. Maggiore è la distanza del centroide dei cluster dal centro, maggiore è la probabilità che si tratti di opinion leader. Nei grafici, gli assi x e y rappresentano rispettivamente il numero di retweet e il numero di like, poiché questi sono i parametri più significativi (poiché erano quelli con i valori più alti) rispetto a quote e reply in questa determinata analisi.

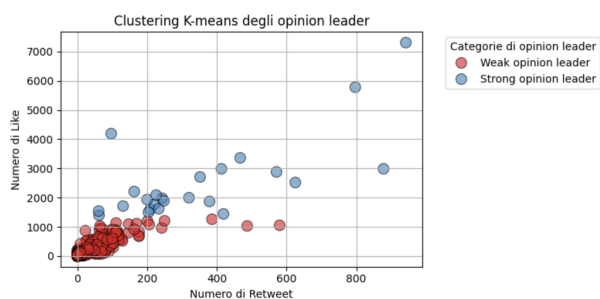
## Risultati analisi opinion leader



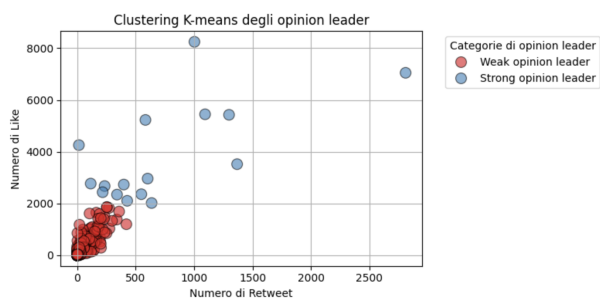
(a) Grafico k-means del picco positivo con token 6563.



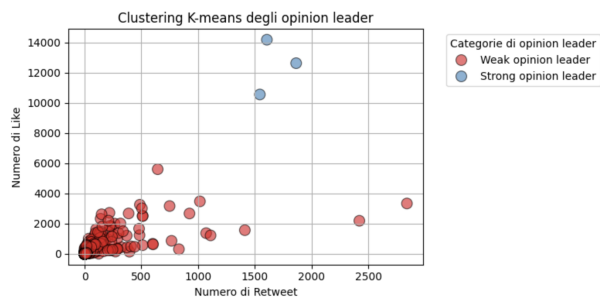
(b) Grafico k-means del picco negativo con token 4037.



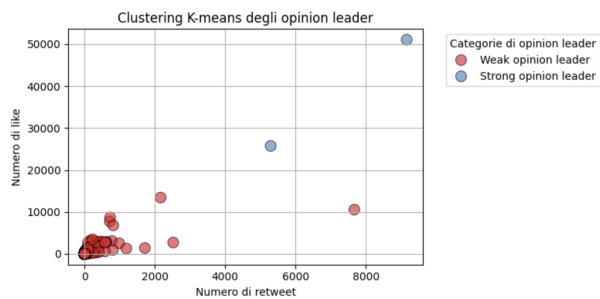
(c) Grafico k-means del picco positivo con token 12247.



(d) Grafico k-means del picco positivo con token 8898.



(e) Grafico k-means del picco positivo con token 4704.



(f) Grafico k-means del picco negativo con token 9032.

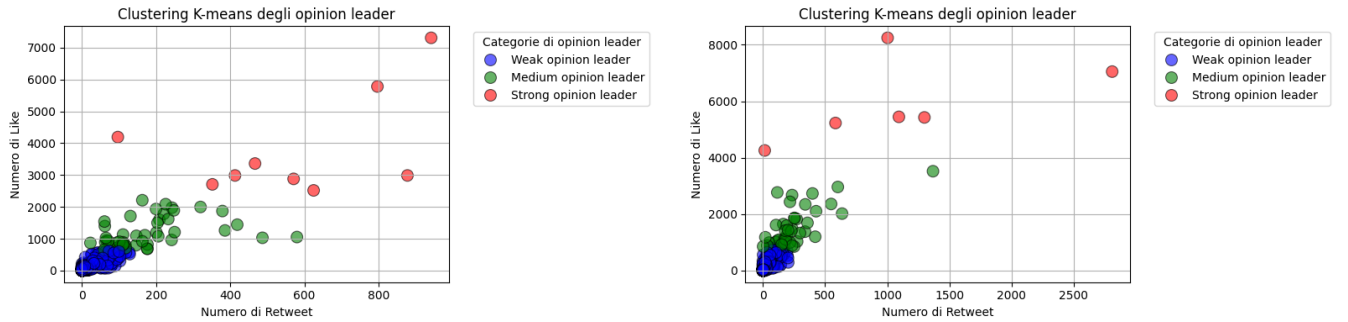
Figura 10: Grafici k-means relativi ai picchi positivi e negativi dei tweets analizzati.

Token ID	Strong opinion leader	Collection
6563	7	doodles
4037	1	doodles
12247	24	doodles, cryptopunks
8898	16	doodles
4704	3	doodles
9032	2	doodles

Tabella 6: Tabella opinion leaders per ogni token.

Dopo aver identificato gli strong opinion leader per ogni picco, abbiamo analizzato i tweet associati e raccolto le loro informazioni generiche, come il nome utente, la descrizione del profilo e la collezione a cui si riferivano i loro tweets, mergiando i dataframes con il dataset del file *"twitter\_tweets\_collection.csv"* e *"twitter\_users.csv"*.

Abbiamo osservato dai grafici e dalla tabella precedenti che ci sono troppi opinion leader forti per i token 12247 e 8898. Di conseguenza, abbiamo deciso di **forzare il valore di k=3** per introdurre un ulteriore cluster (**Medium opinion leader**) al fine di valutare eventuali cambiamenti nell'analisi degli opinion leader.



(a) k-means con k=3 del token 12247

(b) k-means con k=3 del token 8898

Figura 11: Distribuzioni relative ai picchi negativi dei tweet analizzati.

In questo modo i due token avrebbero avuto gli opinion leader distribuiti come nella tabella che segue:

Token ID	Strong opinion leader	Medium opinion leader	Weak opinion leader	Collection
<i>12247</i>	9	49	915	doodles, cryptopunks
<i>8898</i>	6	44	1081	doodles

Tuttavia, **abbiamo optato per utilizzare il valore di k calcolato tramite il metodo della silhouette**, ritenuto più analitico.

Per verificare se gli opinion leader sono effettivamente personaggi influenti nel discorso dei non-fungible token (NFT), abbiamo applicato la sentiment analysis non solo ai contenuti dei tweet, ma anche alle descrizioni delle bio degli opinion leader, come descritto nella sezione 3.3. Dai risultati, abbiamo riscontrato che nessuno degli opinion leader associati ai vari picchi presenta una bio pertinente agli NFT. Solo un utente, un opinion leader di rilievo per il token *12247*, ha una bio rilevante per il contesto, sebbene con un valore di sentiment negativo di -0.1666.

tweet_id	username	retweet_count	like_count	quote_count	reply_count
1364553292991385602	hornedbride	419	1441	7	7

description
elijah • black ndn • 18 • werewolf enthusiast • i draw horror and bears • if you ask me to do nfts im tweeting your ip address

Tabella 7: Dati del tweet e descrizione dell'autore.

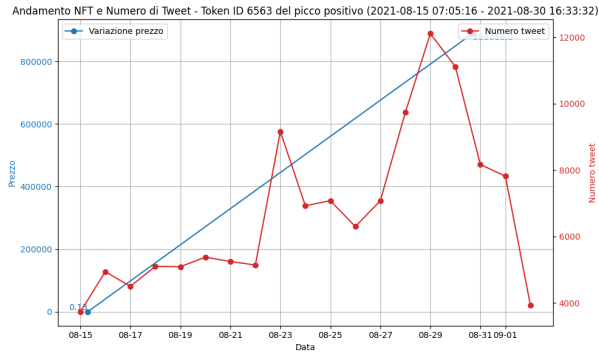
Come mostrato nella tabella 7, questo utente, nonostante abbia una bio pertinente agli NFT, non esercita una grande influenza nella rete, considerando il suo tweet.

### 3.5.4 Analisi influenza NFT sulla quantità di tweets

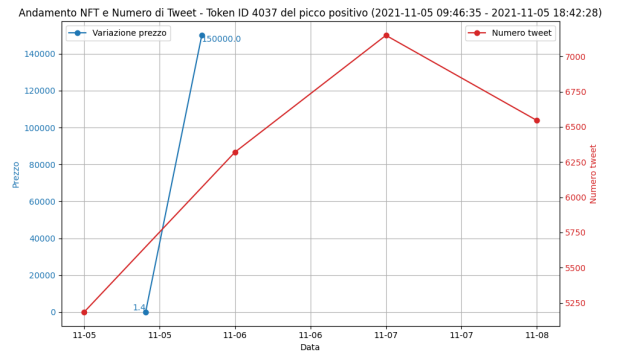
In questa sezione, è stata condotta l'analisi inversa, ovvero lo studio fatto sui top picchi positivi (5) e negativi (1) dei token NFT, concentrandosi sulla variazione di prezzo degli NFT e sulla quantità dei tweet nel periodo di variazione.

Sia per i picchi positivi che per quelli negativi, è stato seguito lo stesso procedimento andando, in primo luogo, ad estrapolare il periodo di inizio e fine della variazione di prezzo del token NFT, selezionando successivamente i tweet pubblicati dal primo giorno della variazione fino ai 3 giorni successivi alla data di fine variazione.

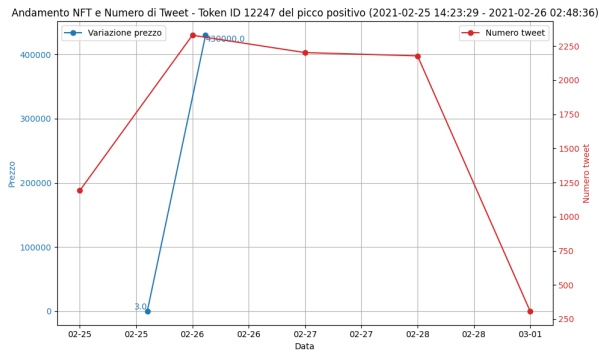
Per ciascun picco, è stato rappresentato graficamente l'**andamento della variazione del prezzo** del token NFT in blu (considerando il prezzo di inizio e di fine) e in rosso l'**andamento della variazione della quantità di tweet** effettuati nello stesso periodo.



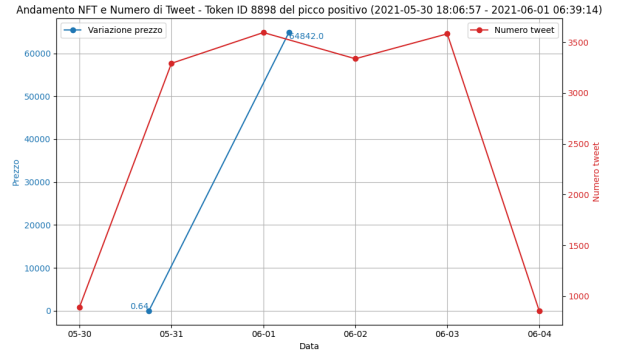
(a) Token ID **6563**.



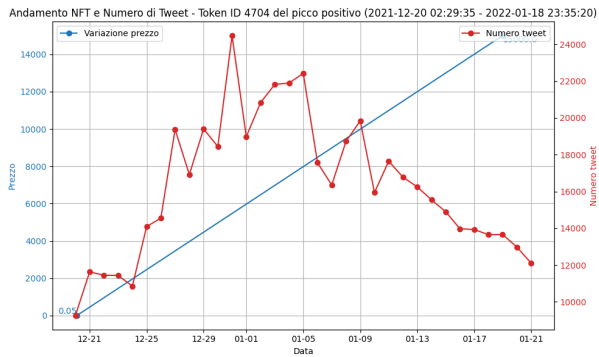
(b) Token ID **4037**.



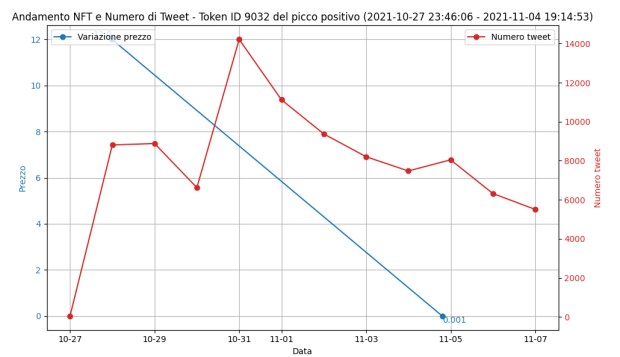
(c) Token ID **12247**.



(d) Token ID **8898**.



(e) Token ID **4704**.



(f) Token ID negativo **9032**.

Figura 12: Grafici per confrontare l'andamento della variazione del prezzo degli NFT e l'andamento della variazione del numero di tweet per ciascuno dei picchi considerati.

## Correlazione di Pearson e p-value

Infine abbiamo calcolato la **correlazione di Pearson** ( $r$ ) che misura la forza della relazione lineare tra due variabili. Per il nostro studio abbiamo utilizzato la variazione percentuale del prezzo del token in secondi e la quantità di tweets riguardanti quel token in un certo periodo critico.

Un valore di  $r$  prossimo a 1 indica una forte correlazione positiva, mentre un valore prossimo a -1 indica una forte correlazione negativa. Un valore prossimo a 0 indica l'assenza di correlazione lineare. Il **valore p** (*p-value*) indica la probabilità che la correlazione osservata sia dovuta al caso. Un *p-value* inferiore a 0.05 generalmente suggerisce che la correlazione è statisticamente significativa.

L'operazione di calcolo della correlazione e del p-value è stata svolta solamente per quei picchi positivi dove c'era un evidente e sensibile correlazione tra un fenomeno (nel nostro caso dei tweets influenti) e la variazione del prezzo. Per svolgere il calcolo della correlazione definito come:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}$$

in cui teniamo conto delle due variabili da esaminare ( $X$  e  $Y$ ), dove  $X_i$  e  $Y_i$  sono i valori delle variabili per l'osservazione  $i$ , e  $\bar{X}$  e  $\bar{Y}$  sono le loro medie, rispettivamente.

La correlazione generale dovrebbe essere calcolata considerando tutte le righe come un insieme di coppie di valori per le due variabili di interesse: '**variazione percentuale per secondo**' e '**numero di tweet**'.

Abbiamo scelto `variazione_per_sec` perché ci interessa analizzare la correlazione immediata tra il numero di tweets e la variazione del prezzo dei token NFT, evidenziando la velocità con cui il mercato reagisce agli stimoli provenienti dai social media nel breve termine.

Durante l'analisi dei **picchi con variazione negativa**, abbiamo identificato **solo un token** il cui prezzo sembrava essere effettivamente influenzato dal fenomeno. A causa di ciò, non siamo riusciti a calcolare la correlazione e quindi il p-value, poiché sono necessari almeno due elementi nel dataset per condurre tali calcoli.



data_inizio	data_fine	token_id	price_trend	trend_change
2021-08-15 07:05:16	2021-08-30 16:33:32	<i>6563</i>	positivo	136359
2021-11-05 09:46:35	2021-11-05 18:42:28	<i>4037</i>	positivo	91506
2021-02-25 14:23:29	2021-02-26 02:48:36	<i>12247</i>	positivo	16568
2021-05-30 18:06:57	2021-06-01 06:39:14	<i>8898</i>	positivo	174162
2021-12-20 02:29:35	2022-01-18 23:35:20	<i>4704</i>	positivo	103908
prezzo_inizio	prezzo_fine	percentuale_variazione	durata	
0.130	888888.0	6.837599e+08	1330096.0	
1.400	150000.0	1.071419e+07	32153.0	
3.000	430000.0	1.433323e+07	44707.0	
0.640	64842.0	1.013146e+07	131537.0	
0.050	15000.0	2.999990e+07	2581545.0	
variazione_per_sec	numero_tweets	corr_pearson	p-value	
514.068082	128501	<b>0.997369</b>	<b>0.046189</b>	
333.225071	25202	<b>0.757791</b>	<b>0.242209</b>	
320.603783	8208	<b>-0.513250</b>	<b>0.376458</b>	
77.023670	15556	<b>-0.698739</b>	<b>0.301261</b>	
11.620909	537653	<b>-0.672743</b>	<b>0.530232</b>	

Correlazione di Pearson (intero dataset): -0.5132495

Valore p (intero dataset): 0.3764577

Figura 13: Correlazioni di Pearson e p-value per ogni token NFT con variazione positiva

- *Token 6563*: La correlazione di Pearson è 0.997, indicando una correlazione positiva quasi perfetta tra la variazione percentuale per secondo e il numero di tweet. Il p-value di 0.046 suggerisce che questa correlazione è statisticamente significativa, con una probabilità molto bassa che sia dovuta al caso.
- *Token 4037*: La correlazione di Pearson è 0.758, indicando una forte correlazione positiva. Tuttavia, il p-value di 0.242 non è sufficientemente basso per affermare che la correlazione sia statisticamente significativa.
- *Token 12247*: La correlazione di Pearson è -0.513, indicando una moderata correlazione negativa. Il p-value di 0.376 non è sufficientemente basso per affermare che la correlazione sia statisticamente significativa.
- *Token 8898*: La correlazione di Pearson è -0.699, indicando una moderata correlazione negativa. Il p-value di 0.301 è alto, suggerendo che la correlazione non è statisticamente significativa.
- *Token 4704*: La correlazione di Pearson è -0.673, indicando una moderata correlazione negativa. Il p-value di 0.530 è alto, suggerendo che la correlazione non è statisticamente significativa.

Considerando **tutti i dati**, i risultati hanno rivelato una correlazione negativa moderata, con un coefficiente di Pearson di **-0.513**. La correlazione quindi può essere descritta come "moderata", il che significa che la relazione tra le due variabili esiste, ma non è estremamente forte. Questo tipo di correlazione può essere dovuto a diverse ragioni, come l'influenza di altri fattori non considerati nell'analisi o la natura complessa delle interazioni tra il prezzo degli NFT e l'attività sui social media.

Tuttavia, il p-value associato a questa correlazione è stato calcolato a **0.376**, indicando che la relazione osservata potrebbe essere attribuibile al caso anziché a una connessione effettiva tra le

variabili. Pertanto, non ci sono evidenze sufficienti per confermare un legame statistico significativo tra l'andamento del prezzo dell'NFT e l'attività sui social media, rappresentata dai tweet.

## 4 Conclusioni

L'analisi integrata dell'andamento del prezzo degli NFT sul mercato e dei post pubblicati dagli utenti su Twitter offre una prospettiva approfondita su come il sentiment espresso sui social media possa influenzare i mercati come quello degli NFT.

Il nostro lavoro è disponibile nella repository pubblica di **GitHub** [2].

### 4.1 Risultati ottenuti

#### 4.1.1 Risultati: se i tweet hanno influenzato gli NFT

Analizzando le tabelle per i tweet dei picchi positivi 3 e per i tweet dei picchi negativi 5 notiamo che:

- Per i **picchi positivi**: possiamo dire che molto probabilmente sono stati influenzati dai tweet, in quanto in tutti i picchi il sentiment positivo è maggiore rispetto a quello neutrale e negativo.
- Per il **picco negativo**: esaminando i tweet pubblicati durante il periodo corrispondente, è possibile notare che il sentiment di questi è in maggioranza positivo e neutrale, mentre i tweet con sentiment negativo sono piuttosto bassi a confronto. Potremmo presumere che il picco negativo del NFT con token 9032 è stato probabilmente influenzato da altri fattori che non riguardano i tweet degli utenti presi in esame.

Ricordiamo, inoltre, che la sentiment analysis ha dei limiti sulla piattaforma Twitter, questo perché i tweets rientrano in un contesto limitato e i testi pubblicati dagli utenti possono contenere ironia e sarcasmo difficili da interpretare.

Per individuare i tweets che hanno influenzato gli NFT, siccome la mole di tweets era ancora molto elevata, abbiamo analizzato gli opinion leader spiegata al Capitolo 3.5.2.

Abbiamo individuato gli strong opinion leader e, per ognuno dei picchi, i relativi tweets che hanno pubblicato. Andremo ora ad analizzare questi tweet.

Token ID	Numero strong opinion leader	Numero tweet con			Collection
		sentiment positivo	sentiment neutrale	sentiment negativo	
6563	7	4	3	0	doodles
4037	1	0	1	0	doodles
12247	24	5	16	3	doodles, cryptopunks
8898	16	2	12	2	doodles
4704	3	2	1	0	doodles
9032	2	0	2	0	doodles

Tabella 8: In questa tabella sono rappresentati i primi 5 picchi positivi dei token NFT e l’ultimo picco negativo (token 9032). Per ciascun token, sono indicati il numero di opinion leader di rilievo, il numero di tweets suddiviso per sentiment e la collezione dei tweets.

In conclusione, abbiamo notato che, oltre al sentiment positivo, è spesso presente anche quello neutrale: i tweet neutrali possono essere comunque **presi in considerazione** poiché non ne parlano male, perciò non influiscono negativamente sull’andamento ma, al contrario, possono **aumentarne l’hype** e la divulgazione. Il **sentiment negativo** è decisamente meno presente e viene sovrastato dal numero dei tweet positivi e neutrali. Per questi motivi, possiamo ritenere che i tweet abbiano **influenzato** l’andamento degli NFT.

Per quanto riguarda il **picco negativo**, non possiamo attribuire l’andamento negativo al contenuto dei tweets in quanto non abbiamo riscontrato tweet con sentiment negativo; erano principalmente neutrali e, leggendo il contenuto dei due tweet, non sono state riscontrate espressioni sfavorevoli nei confronti degli NFT.

#### 4.1.2 Risultati: se gli NFT hanno influenzato i tweet

Grazie ai grafici della figura 12, possiamo analizzare l’andamento del prezzo (rappresentato dalla linea blu che unisce il prezzo iniziale e quello finale) insieme al numero di tweets (rappresentati dalla

linea rossa) pubblicati nello stesso intervallo di tempo, analizzando anche i tre giorni successivi alla data finale della variazione.

Notiamo che, per i picchi positivi, i tweet pubblicati aumentano con l'aumentare della variazione del prezzo degli NFT, suggerendo una correlazione tra queste due variabili.

Sebbene a livello grafico sembri esserci una forte correlazione tra l'andamento del prezzo dei token NFT e la quantità di tweets su di essi, il calcolo della **correlazione di Pearson** ha rivelato un valore di -0.51325, indicando un collegamento statistico non evidente. Il valore di causalità, espresso dal **p-value** di 0.3764577, suggerisce che altri fattori potrebbero influenzare in modo più significativo la dinamica del mercato degli NFT rispetto alla semplice interazione sui social media.

Per il picco negativo, il numero di tweet che seguono l'inizio della sua variazione discendente sembra essere comunque elevato a livello grafico, sebbene non è stato possibile calcolarlo statisticamente.

Questo potrebbe indicare una reazione significativa della comunità (che ne aumenta la propagazione) rispetto al movimento negativo del prezzo degli NFT, sebbene non siano stati rilevati tweet con sentiment positivo, negativo o neutrali.

## 4.2 Limitazioni

Le principali limitazioni di questa analisi sono le seguenti:

- **Identificazione degli opinion leader:** in presenza di un elevato numero di opinion leader per due token specifici, anche se il valore del metodo silhouette suggerisce di utilizzare il k-means per creare due cluster, ciò implica che non esiste un opinion leader predominante in grado di influenzare significativamente la massa degli utenti, e di conseguenza il mercato e il prezzo.
- **Sentiment analysis:** la mancata precisione di questo metodo (come il limitato contesto dei tweet e l'utilizzo di sarcasmo e ironia) ha fatto sì che certi risultati fossero dubbi e poco chiari, complicando così l'analisi.

- **Dataset troppo vasto:** la enorme mole di dati ci ha spesso impedito di effettuare le diverse analisi su un periodo di tempo più vasto e una quantità di dati maggiori. Per questo motivo, siamo stati sempre costretti a dover filtrare i dati, svolgendo prima delle analisi per individuare quelli più futili e successivamente rimuoverli. Inoltre, sempre per lo stesso motivo, abbiamo spesso dovuto usufruire di dataframes per esportare i dati in formato .csv e importare questi ultimi in un nuovo file per non eseguire di nuovo tutto il codice dovuto alle tempistiche di esecuzione molto elevate.

### 4.3 Lavori futuri

Per approfondire e migliorare lo studio svolto, proponiamo i seguenti sviluppi futuri:

- **Analisi più dettagliata del sentiment.** Implementare metodi di analisi del sentiment più avanzati aiuterebbero a rilevare il sentiment con maggiore precisione. La classificazione dei tweet potrebbe essere fatta in categorie più specifiche. Anziché limitarsi a sentimenti positivi, negativi e neutrali, una categorizzazione più dettagliata potrebbe offrire più sfumature sulle emozioni degli utenti.
- **Identificazione dei nodi influenti nella rete di Twitter.** Attraverso un'analisi della rete sociale degli utenti twitter e dei tweet è possibile utilizzare metriche di centralità per scoprire quali utenti o quali tweet hanno un maggiore impatto sulle discussioni.
- **Analisi degli impatti esterni.** Esaminare come eventi esterni possano influenzare i risultati delle analisi potrebbe aiutare a studiare i picchi del mercato. Nel nostro progetto, non sempre i risultati sono stati chiari, e ciò potrebbe essere dovuto a eventi esterni che hanno avuto un impatto sul volume e sul contenuto dei tweet. Alcuni esempi di fenomeni da studiare potrebbero includere eventi di mercato più ampi (regolamentazioni, lanci di nuovi prodotti, ecc...) oppure riguardare la politica e altri mondi.
- **Perfezionamento della ricerca sugli opinion leader.** Per comprendere chi sono gli opinion leader più influenti e capire come le loro opinioni impattano le variazioni di prezzo degli

NFT, è possibile attraverso un'analisi più sofisticata degli opinion leader e il loro impatto immediato sul mercato. Un esempio, potrebbe essere analizzare i follower degli utenti andando, però, a distinguere quelli reali e quelli acquistati.

## 5 Lavori correlati

Per svolgere questo studio, abbiamo usato come linea guida quella seguita dal lavoro **"On Using Twitter to Understand the Stablecoin Terra Collapse"** [3] di Stefano Ferretti e Marco Furini.

Questo articolo analizza la criptovaluta "Terra" che permette la creazione di determinate stablecoin ed esamina le opinioni e i pensieri delle persone sull'argomento sfruttando i dati di Twitter. Per quanto riguarda la piattaforma, ne individua e valuta l'influenza degli opinion leader.

Questo articolo è tornato molto utile al nostro studio per capire come analizzare gli utenti di Twitter e, nello specifico, come estrapolare i tweet virali e distinguere gli opinion leader.



# Bibliografia

- [1] Il nuovo sistema per citare i tweet su twitter, Giugno 2014.
- [2] Repository del progetto su github, <https://github.com/riccardospini98/Blockchain>, 2024.
- [3] Stefano Ferretti e Marco Furini. On using twitter to understand the stablecoin terra collapse, September 2023.
- [4] Alex McFarland. 10 migliori librerie python per l'analisi del sentiment, January 2024.
- [5] Fastweb plus. Cos'è twitter, come funziona e cosa si può fare.