



**UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”**

Corso di Laurea Magistrale o Specialistica in Data Science per l'Economia e le Imprese

**Rilevazione e analisi degli eventi di pump and dump nel mercato
delle criptovalute**

**Detection and analysis of pump and dump events in the
cryptocurrency market**

Relatore: Chiar.mo
Prof. Luca Virgili

Tesi di Laurea di:
Dumitru Razvan
George

Anno Accademico 2022 – 2023

Sommario

Introduzione	8
Blockchain e criptovalute	12
Panoramica sulla tecnologia blockchain.....	12
Principi alla base della blockchain	15
Componenti della blockchain.....	15
Funzionamento della Blockchain	17
Il mercato delle criptovalute	19
Le origini delle criptovalute	19
Il fenomeno delle memecoin.....	21
Descrizione degli eventi	24
Il Pump and Dump nel mercato delle Criptovalute.....	26
Lavori correlati al fenomeno.....	33
Rilevazione ed analisi degli eventi Pump and Dump.....	37
Panoramica del dataset	37
Eventi di Pump and Dump	37
Transazioni.....	43
Definizione delle feature per la classificazione	53
Feature in letteratura	53
Matrix Profile	57
Classificazione degli eventi di Pump and Dump	62
Discussione	67
Conclusione	69
Bibliografia	71
Sitografia.....	72
Ringraziamenti	74

Indice delle figure

Figura 1 – Funzionamento della tecnologia blockchain nel mondo delle criptovalute	17
Figura 2 – Andamento del titolo Gamestop durante i giorni dell'intervento da parte della comunità di Reddit	23
Figura 3 – Andamento del prezzo della criptovaluta AppCoins e il numero di unità scambiate, prima, dopo e durante l'evento	30
Figura 4 – Frammento di conversazione ripresa dal gruppo Telegram in cui viene annunciato un evento Pump and Dump.....	32
Figura 5 – Analisi dei gruppi di Pump and Dump trovati su Telegram e Discord	39
Figura 6 – Numero di eventi Pump and Dump individuati omettendo quelli promossi da più gruppi	40
Figura 7 – Variazioni di prezzo della Bored Coin durante i Pump and Dump del 2018	44
Figura 8 – Numero di unità scambiate durante i Pump and Dump del 2018.....	46
Figura 9 – Fluttazioni del prezzo della Bored Coin durante i Pump and Dump del 2019.....	48
Figura 10 – Andamento storico del prezzo del Bitcoin nel corso degli anni (2015 – 2024).....	49
Figura 11 – Numero di unità scambiate durante i Pump and Dump del 2019.....	49
Figura 12 – Volumi di scambio in Bitcoin della Bored Coin nei Pump and Dump del 2019	50
Figura 13 – Fluttazioni del prezzo della Bored Coin durante i Pump and Dump del 2020.....	51
Figura 14 – Numero di unità scambiate durante i Pump and Dump del 2020.....	52
Figura 15 – Andamento del prezzo prima, durante e dopo l'evento.....	53

Figura 16 – Numero di unità scambiate nei momenti prima, durante e dopo l'evento	54
Figura 17 – Confronto delle distanze tra le sottosequenze di una serie storica di lunghezza $n=13$	59
Figura 18 – Rilevazione dei motifs su un grafico delle distanze	61

Elenco delle tabelle

Tabella 1 - Analisi dei gruppi di Pump and Dump trovati su Telegram e Discord	28
Tabella 2 - Numero di eventi Pump and Dump organizzati da ciascun gruppo e numero di criptovalute.....	42
Tabella 3 – Performance dei due modelli di partenza.....	63
Tabella 4 – Performance degli esperimenti condotti per individuare la corretta window size	64
Tabella 5 – Performance ottenute con le feature calcolate con il Matrix Profile ..	65

Introduzione

Con il termine *Pump & Dump* ci si riferisce ad una tipologia di frode che consiste nel far lievitare artificialmente il prezzo di un'azione a bassa capitalizzazione con l'obiettivo finale di vendere titoli azionari acquistati a buon mercato ad un prezzo superiore.

Sul mercato azionario questi comportamenti sono stati definiti illegali e grazie ai numerosi controlli degli organi di vigilanza sui mercati sono spesso difficili da attuare.

La Securities and Exchange Commission (SEC, l'organo di vigilanza americano in ambito finanziario) intenta regolarmente cause civili per questa tecnica fraudolenta quando è applicata ai titoli quotati in borsa. Le manipolazioni delle criptovalute non sono diverse, ma le autorità di regolamentazione non hanno ancora deciso come intervenire in questo mercato come riportato anche in un articolo tratto da Milano Finanza¹.

Questa lacuna normativa ha permesso la nascita di comportamenti opportunistici nel mercato dei prodotti finanziari digitali come le criptovalute e gli NFT.

L'analisi compiuta dalla Federal Trade Commission (FTC, un'agenzia governativa americana che si occupa di promuovere la tutela dei consumatori e

¹Come si manipolano le criptovalute

<https://www.milanofinanza.it/news/pump-and-dump-cosi-si-manipolano-le-criptovalute-201808061936404179>

l'eliminazione e la prevenzione di pratiche commerciali anticoncorrenziali) ha rilevato che le truffe in questo ambito sono preoccupanti dato che circa 7.000 persone hanno perso più di 80 milioni di dollari in truffe di criptovalute da ottobre 2020 a marzo 2021, questo è un dato rilevante se comparato alle 570 truffe di investimento in criptovalute e ai 7,5 milioni di dollari di perdite durante gli stessi mesi dell'anno precedente.²

La società britannica di analisi Elliptic, che dal 2013 si occupa di prevenire la criminalità finanziaria nelle criptovalute, ha analizzato oltre 80 truffe, tra le più significative segnalate sui social media da luglio 2021, per un totale di 17.789 NFT rubati e un valore complessivo di oltre 100 milioni di dollari.³

Un altro fattore che incoraggia la diffusione di questi fenomeni è la facilità, di raggiungere via Internet un numero sempre più elevato di utenti in pochi secondi.

Nel caso che andremo ad analizzare nello specifico la truffa viene effettuata tramite l'app gratuita di messaggistica *Telegram*. Quest'app infatti permette la creazione di gruppi molto numerosi detti anche "canali", nei quali è possibile diffondere informazioni ad un'ampia platea di utenti.

²Truffe effettuate sulle criptovalute

<https://www.cybersecurity360.it/nuove-minacce/le-frodi-tramite-criptovalute-come-riconoscerle-i-metodi-piu-comuni-e-come-difendersi/>

³Dati sulle truffe ai danni di proprietari di NFT

<https://collezionedatiffany.com/rubati-piu-di-100-milioni-di-in-nft-in-un-solo-anno>

Inoltre, i sistemi di *Pump and Dump* possono apparire come opportunità redditizie, ma in genere sono gli organizzatori principali ad ottenere il massimo profitto. Ciò è dovuto al fatto che sono consapevoli del momento in cui avverrà l'aumento del prezzo della criptovaluta. È probabile che alcuni partecipanti all'evento ricoprano livelli diversi e ricevano un guadagno in base al ruolo che ricoprono nell'organizzazione dell'evento. La realtà, è che si rischia di perdere denaro, soprattutto se non si ricopre un ruolo importante nell'organizzazione dell'evento.⁴

L'obiettivo di questo elaborato è quello di individuare gli eventi *Pump and Dump*, partendo da un dataset costruito con l'estrazione e l'analisi delle conversazioni su Telegram, Reddit e Twitter di 20 gruppi differenti dai quali sono stati identificati individuati 904 diversi eventi.

In seguito, con lo scopo di identificare gli eventi di *Pump and Dump* verranno utilizzati due algoritmi di classificazione (Random Forest e XGBoost), come fatto nel paper di La Morgia et al. [10].

Infine, verranno testate le feature estratte dal Matrix Profile, una tecnica di analisi delle serie storiche che permette di individuare pattern anomali nascosti nei dati.

La tesi è organizzata come segue: nel secondo capitolo sarà introdotta la “blockchain”, cioè la tecnologia che sta alla base del mondo delle criptovalute e

⁴Il funzionamento degli schemi Pump and Dump.

<https://coinrivet.com/it/guides/security/what-is-a-pump-and-dump-crypto-scheme/>

che ne regola gli scambi e verrà introdotto il concetto di criptovaluta, cercando di capire da cosa nasce la sua recente popolarità.

Nel terzo capitolo verranno descritti gli eventi *Pump and Dump* e si vedrà come il fenomeno non è nuovo nella storia dei mercati finanziari ma si è solamente evoluto e verrà fatta una panoramica sui lavori che hanno posto un faro d'attenzione a questo fenomeno.

Nel quarto capitolo invece verrà fatta un'analisi descrittiva per poi focalizzarsi sull'approccio per l'individuazione degli eventi di *Pump and Dump* all'interno del dataset. Verranno spiegate le feature utilizzate, gli algoritmi proposti e l'analisi.

Nel quinto capitolo, invece, verranno discussi i risultati ottenuti e infine nel sesto capitolo verranno ripercorsi i passi dell'analisi fatti e saranno presentati possibili lavori futuri.

Blockchain e criptovalute

Panoramica sulla tecnologia blockchain

La blockchain è una struttura dati che consiste in elenchi crescenti di record, denominati "blocchi", collegati tra loro in modo sicuro utilizzando la crittografia. Ogni blocco contiene un hash crittografico del blocco precedente, un *timestamp* e dati di transazione. Poiché ogni blocco contiene informazioni sul blocco precedente, questi formano effettivamente una catena con ogni blocco aggiuntivo che si collega a quelli precedenti. Di conseguenza, le transazioni blockchain sono irreversibili in quanto, una volta registrate, i dati in un determinato blocco non possono essere modificati retroattivamente senza alterare tutti i blocchi successivi.

La blockchain rientra nella più ampia famiglia dei registri distribuiti (*distributed ledger*), ossia sistemi che si basano su un registro replicato, condiviso e sincronizzato tra più soggetti presenti in molteplici luoghi, ma comunque appartenenti alla medesima entità. Nel caso della *Blockchain* non è richiesto che i nodi coinvolti conoscano l'identità reciproca o si fidino l'uno dell'altro perché, per garantire la coerenza tra le varie copie, l'aggiunta di un nuovo blocco è globalmente regolata da un protocollo condiviso. Una volta autorizzata l'aggiunta del nuovo blocco, ogni nodo aggiorna la propria copia privata. La natura stessa della struttura dati garantisce l'assenza di una sua manipolazione futura.

Le caratteristiche che accomunano i sistemi sviluppati con le tecnologie della blockchain e dei registri distribuiti sono: digitalizzazione dei dati, decentralizzazione, disintermediazione, tracciabilità dei trasferimenti, trasparenza/verificabilità, immutabilità del registro e programmabilità dei trasferimenti. Grazie a tali caratteristiche, la blockchain è considerata pertanto un'alternativa in termini di sicurezza, affidabilità, trasparenza e costi alle banche dati e ai registri gestiti in maniera centralizzata da autorità riconosciute e regolamentate (pubbliche amministrazioni, banche, assicurazioni, intermediari di pagamento).⁵

Vediamo alcune applicazioni della tecnologia blockchain:

Nel settore delle utilities le compagnie energetiche utilizzano la tecnologia blockchain per creare piattaforme di commercio energetico peer-to-peer e semplificare l'accesso all'energia rinnovabile. I proprietari di case con pannelli solari utilizzano questa piattaforma per vendere la propria energia solare in eccesso ai vicini. Vengono anche organizzate delle raccolte di fondi (in inglese *crowdfunding*) per sponsorizzare l'acquisto di pannelli solari in comunità che non hanno accesso all'energia per poi ricavarne un compenso una volta costruiti.

Nel mondo dell'intrattenimento (musica e film) la tecnologia viene utilizzata per gestire i dati del diritto d'autore in modo da garantire il giusto compenso agli

⁵Definizione di blockchain.

<https://it.wikipedia.org/wiki/Blockchain>

artisti. A questo scopo è nata la startup “Corite” che coinvolgendo attivamente i fan. Dalla piattaforma, gli artisti lanciano le loro mission: un crowdfunding su un loro singolo, sui cui i fan investono. Le cifre sono volutamente basse – 10 dollari a testa per coinvolgere la più ampia *fanbase*, evitando la concentrazione su un singolo finanziatore. Questo permette agli artisti di ottenere una base economica, mantenendo il controllo creativo. Al tempo stesso i fan possono rientrare del loro investimento, partecipando alla creazione e alla promozione, con future royalties. Ultimo, ma non per importanza, è il settore finanziario dove la blockchain viene utilizzata per gestire pagamenti online, trading e conti.⁶

In questo senso sta agendo JP Morgan (un importante istituto di credito americano) che sta iniziando una partnership con sei banche indiane per velocizzare la modalità di liquidazione delle operazioni di borsa tra le banche; questo nuovo sistema oltre a velocizzare le operazioni le renderà anche disponibili 24 ore su 24, sette giorni su sette.

⁶Le principali applicazioni della tecnologia blockchain.

<https://aws.amazon.com/it/what-is/blockchain/>

Principi alla base della blockchain

- **Decentralizzazione:** poiché il controllo e il potere decisionale sono trasferiti da un singolo individuo o un'organizzazione agli individui della rete.
- **Immutabilità:** questo implica il fatto che i partecipanti non possono alterare in alcun modo le informazioni presenti nella blockchain. Una volta che una transazione viene registrata viene trascritta su un libro mastro condiviso e se questa contiene un errore è necessario registrarne una nuova per eliminare la vecchia ma entrambe vengono registrate sul libro mastro.
- **Consenso:** questo principio afferma che è possibile registrare nuove transazioni ogni qualvolta la maggioranza dei partecipanti alla blockchain presta il consenso per farlo.

Componenti della blockchain

- 1) **Nodi:** sono le macchine su cui è installato l'applicativo che consente di verificare lo stato dei registri, dei blocchi e delle transazioni in essi contenuti. La rete dei nodi è di natura peer-to-peer; pertanto, chiunque può contribuire alla sua implementazione, godendo degli stessi diritti di tutti gli altri, mentre la gestione è demandata ad un applicativo open source che gestisce ogni aspetto del funzionamento della Blockchain;

- 2) **Transazione:** è costituita dai dati relativi ad uno scambio di valuta, alla stipulazione/deposito di un contratto ed altri accordi, a prescindere alla loro valenza giuridica. Gli estremi della transazione vengono validati con strumenti crittografici ed archiviati nei blocchi;
- 3) **Blocco:** contengono le transazioni validate e sono concatenati per via di un valore crittografico che lega un singolo blocco a quello successivo, formando appunto una catena di blocchi (Blockchain);
- 4) **Ledger:** registro pubblico che contiene i blocchi validati, il cui database è distribuito a tutti i nodi della Blockchain. Una volta inseriti ed approvati dalla maggioranza dei nodi, sulla base del consenso distribuito, i blocchi diventano immutabili a livello cronologico. Questo particolare consente di evitare il double spending, ossia la possibilità di transare lo stesso bene a più soggetti;
- 5) **Hash:** è un'operazione basata sulla crittografia asimmetrica, che consente di codificare una stringa di testo/numeri in maniera irreversibile, in modo che qualsiasi nodo possa verificare la sua validità, senza poter leggere espressamente il contenuto della transazione. L'hash viene impiegato in vari momenti della costituzione della catena e restituisce un codice che rende di fatto univoco ogni elemento previsto. Tra i principali algoritmi crittografici utilizzati ritroviamo il SHA-256.

Che cos'è una blockchain e come funziona?

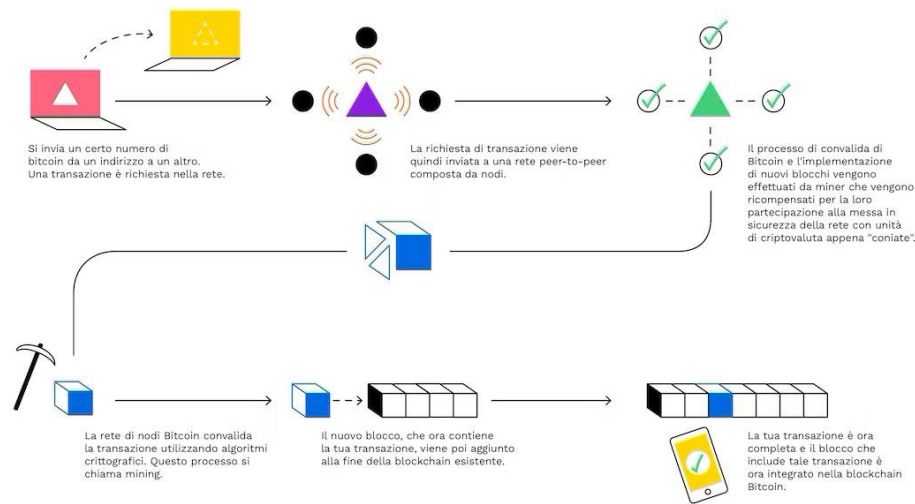


Figura 1 – Funzionamento della tecnologia blockchain nel mondo delle criptovalute

Funzionamento della Blockchain

Com'è mostrato anche in Figura 1, il modello utilizza una combinazione di una firma digitale e una marca temporale, nota anche come *timestamp*. La prima assicura che sia il mittente che il destinatario (o meglio, i loro indirizzi digitali ID) di qualsiasi tipo di messaggio, inclusa una transazione finanziaria, siano identificati in modo affidabile.

Il secondo consente che un insieme di messaggi venga comunicato e scritto nel registro di tutti gli altri nodi della rete e reso irreversibile dopo che è stato validato con una marca temporale da un nodo selezionato casualmente da un modello matematico robusto.

Si tratta di un vero e proprio libro contabile accessibile a tutti, un registro che si aggiorna automaticamente e identicamente su ogni nodo della rete. Gli unici nodi utilizzano software di crittografia per verificare tutte le operazioni; questo software verifica un pacchetto di dati con una chiave privata, che viene utilizzata per firmare le transazioni e che è disponibile per essere controllata in qualsiasi momento.

Tutto il sistema protegge l'identità digitale di chi ha autorizzato gli scambi, definita principalmente come ID, cioè l'indirizzo digitale di un conto che non è necessariamente associato ad un individuo o ad una società.

La caratteristica principale del processo è che il funzionamento non è garantito da un ente centrale; invece, ogni transazione viene validata e approvata dall'interazione di tutti i nodi. Di conseguenza, le transazioni che si verificano all'interno della rete vengono registrate e validate, eliminando la necessità di terze parti "fidate".

Una garanzia condivisa viene creata quando tutti i nodi della rete confermano che una nuova transazione viene inserita nella blockchain. La registrazione di un'operazione in un registro condiviso inoltre permette di conservare nelle transazioni i dati delle parti, l'ammontare, le date di ogni trasferimento di un bene e molto altro.⁷

⁷Il funzionamento della Blockchain.

(<https://www.borsaitaliana.it>)

Il mercato delle criptovalute

Le origini delle criptovalute

La “criptovaluta” è una moneta virtuale, come suggerisce anche il nome è composto da due termini: cripto e valuta.

Il termine “valuta” ne evidenzia il fatto che gli viene attribuito un valore, il termine “cripto” invece denota il fatto che siamo in presenza di qualcosa di criptato al quale è possibile accedere solo se ci conosce la chiave di accesso.

Questo tipo di moneta non esiste in forma cartacea come il denaro che siamo abituati a vedere tradizionalmente tuttavia nel mondo delle monete virtuali vengono ripresi alcuni concetti utilizzati per le monete a corso legale, come ad esempio il “portafoglio” che in questo caso viene definito “portafoglio digitale” o “e-wallet”.

La criptovaluta più nota al momento è il **Bitcoin**, nata dall’idea di Satoshi Nakamoto nel 2008.⁸

L’idea alla base era di creare una forma di moneta digitale che fosse fuori dalla giurisdizione di governi e banche centrali; infatti, le criptovalute non sono state ancora giuridicamente riconosciute e non fanno parte delle monete accettate dalle banche centrali.

⁸Com’è nato il Bitcoin.

<https://www.money.it/Chi-e-Satoshi-Nakamoto-creatore-Bitcoin>

Il sistema di scambio del Bitcoin è regolato dalla blockchain, che, come abbiamo spiegato nel precedente paragrafo, è un registro condiviso e immutabile che permette di registrare tutte le transazioni avvenute: al fine di evitare quasi del tutto comportamenti opportunistici e truffe tra i partecipanti allo scambio.

Il valore del Bitcoin, come quello delle altre criptovalute è basato sul principio della scarsità: non potranno mai essere creati più di 21 milioni di Bitcoin.

A seconda della criptovaluta cambia anche il sistema di blockchain utilizzato, questo può comportare una variazione nei tempi di scrittura della transazione che è di qualche minuto per i Bitcoin ma diventa di qualche secondo se lo scambio viene effettuato in Ethereum⁹.

Dalla loro nascita le criptovalute hanno subito una grande accelerazione, soprattutto negli ultimi anni. Con l'avvento delle informazioni sul loro funzionamento e gli enormi guadagni derivanti da queste hanno permesso la diffusione di una grande varietà di questi prodotti finanziari con caratteristiche diverse, i più diffusi sono:

- **Monete:** tutte le criptovalute che sfruttano una blockchain indipendente, cioè una propria infrastruttura. Tra queste rientrano il Bitcoin e l'Ethereum, quest'ultimo infatti dispone dell'infrastruttura "Ether".

⁹Bitcoin e le altre criptovalute.

(<https://www.geopop.it/bitcoin-e-le-altre-criptovalute-cosa-sono-e-come-facciamo-ad-attribuirgli-un-valore/>)

- **Altcoin:** tutte le monete diverse dal Bitcoin ma che hanno un funzionamento ad esso analogo. Alcune hanno una quantità limitata, come il Bitcoin che è fissato a 21 milioni, altre come il Dogecoin invece offre una quantità illimitata di monete.
- **Token:** sono degli asset digitali che possono essere scambiati o acquistati come le monete digitali, la differenza sta nel fatto che sfruttano l'infrastruttura di una blockchain appartenente ad altri. In gergo vengono definiti "asset non nativi". Un esempio è Tether che sfrutta la tecnologia blockchain di altre monete: Ethereum, Chainlink, Uniswap e Polygon.

Il fenomeno delle memecoin

Le *memecoin*, che sono criptovalute che si ispirano ai *meme* che si trovano su Internet e sui social media, sono nate insieme alle criptovalute più note e "stabili" come Bitcoin e Ethereum.

Dogecoin (DOGE) è stato il primo *memecoin*, è stato lanciato nel 2013 come una parodia del popolare meme "Doge" di un cane giapponese di razza Shiba Inu.

Il prezzo delle *memecoin* ha un andamento molto variabile, sono molto guidate dalla comunità e, grazie alla *community* online e alla FOMO (Fear of missing

out), possono guadagnare popolarità in pochissimo tempo. Tuttavia, quando i trader cercano un'altra *memecoin*, il loro prezzo può scendere inaspettatamente. Queste criptovalute hanno spesso un'offerta enorme o infinita, ad esempio, Shiba Inu (SHIB) ha un'offerta totale di un quadrilione di token, mentre DOGE non ha una quantità massima di token disponibili, e ci sono già più di cento miliardi di token in circolazione. L'enorme offerta di *memecoin* spiega i prezzi relativamente bassi, poiché normalmente non bruciano monete. Puoi acquistare milioni di unità di un token meme con solo 1 dollaro.

Precisiamo che il "burning of coins" è un processo che rimuove definitivamente una criptovaluta dalla circolazione, diminuendo così l'offerta totale.

L'avvento della popolarità delle *memecoin* sembra sia nata durante la pandemia di COVID-19 nel 2020, come strumento di protezione contro l'inflazione da parte degli investitori.

L'episodio che ha sancito la nascita delle *memecoin* è quello di Gamestop.

All'inizio del 2021 la società di hedge-fund Melvin Capital, che scommetteva sul ribasso di Gamestop, ha perso oltre il 53% dei suoi asset a causa del "braccio di ferro" condotto contro gli utenti del *subreddit* 'r/WallStreetBets' che invece scommettevano sul loro rialzo. Questi piccoli investitori, coordinandosi su Reddit sono riusciti a scrivere una pagina importante nei libri di storia finanziaria acquistando azioni del titolo e facendolo aumentare dai 39,91 dollari del 14 gennaio 2021 ai 347,51 dollari del 27 gennaio fino a raggiungere il picco di 460 il

28 dello stesso mese (come mostrato dal grafico dell'andamento del prezzo del titolo ripreso dal "New York Times" nella Figura 2).

Successivamente le piattaforme di trading hanno sospeso la possibilità di acquistare il titolo per risparmiare, da una fine peggiore, gli hedge fund.¹⁰

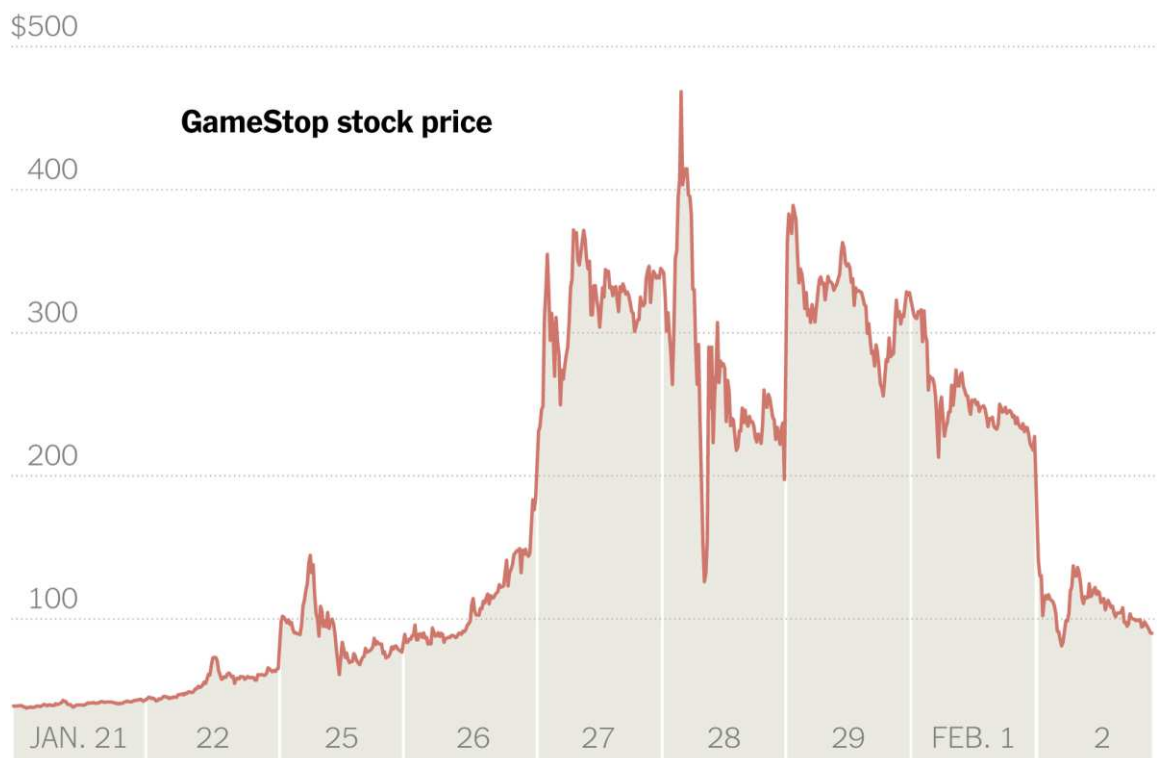


Figura 2 – Andamento del titolo Gamestop durante i giorni dell'intervento da parte della comunità di Reddit

¹⁰Il caso Gamestop.

https://it.wikipedia.org/wiki/GameStop#Caso_boom_azioni_GameStop

Descrizione degli eventi

Il *Pump and Dump* è una truffa che consiste nel rialzo pilotato di un titolo che si possiede attraverso informazioni false o ingannevoli (pump) in modo da rivendere il titolo posseduto ad un prezzo maggiore rispetto a quello di acquisto (dump) e ottenerne un profitto. Nel momento in cui il broker che possiede il titolo inizia a vendere il prezzo inizia a scendere e gli altri investitori ne traggono una perdita. Generalmente questo tipo di truffa è effettuata su titoli a bassa capitalizzazione, per i quali cioè, è possibile acquisire un grande numero di azioni ad un prezzo molto basso ed il loro prezzo è molto più facile da manipolare.

Per attrarre potenziali investitori i broker erano tenuti a contattare le vittime telefonicamente, oggi grazie ad internet i metodi di comunicazione sono aumentati (e-mail, giornali online, siti di investimento) ed è diventato più semplice e veloce coinvolgere una platea più ampia di persone.

Il funzionamento dello schema è molto semplice: il broker contatta gli investitori riferendo loro di essere in possesso di informazioni riservate su un titolo e che li farà guadagnare se investono. Se il broker riesce a guadagnare la fiducia degli investitori ed iniziano ad acquistare, la domanda per l'acquisto del titolo aumenta e altrettanto fa il prezzo, portando la gente a credere che quello sia un investimento sicuro. La notizia che si diffonde fa aumentare il numero di investitori e quando viene raggiunto un determinato livello di prezzo, il broker inizia a vendere le proprie azioni facendo profitti. Al contrario, gli investitori

venderanno ad un prezzo più basso o non riusciranno a vendere i propri titoli in tempo e accumuleranno perdite.

Negli anni 90 è celebre il caso della “Stratton Oakmont”, società che si occupava della compravendita di titoli a bassa capitalizzazione che tramite la diffusione di notizie false o ingannevoli “gonfiavano” i prezzi delle azioni da loro acquistate. Il co-fondatore della società, Jordan Belfort, è stato accusato di aver partecipato allo schema fraudolento e poi condannato a quattro anni di carcere, nonché al risarcimento delle persone truffate. La vicenda ha spinto Belfort a scrivere un libro dal quale è stato ispirata la trama del celebre film “The Wolf of Wall Street”.¹¹

Un altro schema *Pump and Dump* nel mondo della finanza è quello della società energetica americana “Enron” avvenuta nel 2001.

Questa società, considerata tra le più solide in America (la settima multinazionale negli Stati Uniti per importanza), ebbe una crescita esponenziale negli ultimi dieci anni di attività fino a fallire improvvisamente. Le cause che portarono al suo fallimento furono da una parte la dichiarazione di falsi profitti che servivano a far sembrare la società più solida di quanto fosse. A rafforzare questa posizione ci fu lo schema di società-fantasma costituite nei paradisi fiscali che le permettevano di evadere una quota considerevole di tasse facendole registrare profitti maggiori con

¹¹La vicenda della Stratton Oakmont

https://it.wikipedia.org/wiki/Stratton_Oakmont

la compiacenza del mondo politico americano che concesse agevolazioni alla società in cambio di aiuti nelle campagne elettorali o donazioni. Considerando che la società avesse accumulato debiti per 10 miliardi di dollari il suo fallimento è stato piuttosto grave: i dipendenti vennero licenziati in massa e gli azionisti della società persero il denaro investito. Il governo degli Stati Uniti creò una commissione d'inchiesta per indagare sullo schema fraudolento messo in atto, questo portò alla condanna degli amministratori con pene che andarono dai diciotto mesi ai ventiquattro anni.¹²

Il Pump and Dump nel mercato delle Criptovalute

In sostanza realizzare uno schema di *Pump and Dump* nei mercati finanziari tradizionali è molto complesso e spesso, come visto nei casi discussi precedentemente, porta a indagini da parte delle autorità vigilanti e a conseguenze negative per chi li mette in pratica. Nei mercati delle criptovalute, il fatto che non vi sia ancora un insieme di norme che regolamentano questo strumento finanziario, gioca a favore della diffusione del fenomeno.

Josh Kamps et al. [1] cercano di dare una panoramica del fenomeno e dell'evoluzione del *Pump and Dump* negli ultimi anni. Rispetto al mercato delle azioni a bassa capitalizzazione (le cosiddette *penny stocks*) per le criptovalute

¹² Il fallimento di Enron

<https://it.wikipedia.org/wiki/Enron>

sono nati una grande quantità di gruppi su *Telegram*¹³, nelle *chat-room* di Discord¹⁴, o Reddit¹⁵ con lo scopo di organizzare gli eventi e selezionare la criptovaluta da “pompare”.

Il numero dei partecipanti ad un gruppo può oscillare tra qualche migliaio di iscritti a centinaia di migliaia. Nel nostro caso, i gruppi dai quali sono stati estratti gli eventi che andremo ad analizzare contano in alcuni casi diverse decine di migliaia di iscritti e alcuni sono attivi sia su *Telegram* che su *Discord* (come mostrato nell'analisi riportata in Tabella 1), questo permette di ampliare la potenziale platea di partecipanti all'evento.

Un altro fattore che emerge dall'analisi delle chat è che questi gruppi hanno una costruzione gerarchica; infatti, sulla base del numero di utenti che si riescono a portare sul canale viene riconosciuto un merito che permette di essere avvisati dell'inizio dell'evento prima degli altri utenti del canale. Altri gruppi invece prevedono il pagamento di una *subscription* per prendere parte agli eventi prima degli altri mentre altri non richiedono nessuna forma di riconoscimento verso gli organizzatori degli eventi.

¹³ Telegram

<https://it.wikipedia.org/wiki/Telegram>

¹⁴ Discord

<https://it.wikipedia.org/wiki/Discord>

¹⁵ Reddit

<https://it.wikipedia.org/wiki/Reddit>

Nome del gruppo	Iscritti su Telegram (#)	Iscritti su Discord (#)	Gerarchia del gruppo	Piattaforma principale	P&D (#)	Volume di scambi medio (\$)
Big Pump Signal	72.097	104.830	Affiliazione	Binance	32	7.245.437
Trading Crypto Guide	91.725	-	Vip	Binance	17	2.442.923
Crypto Coin B	166.689	-	Vip	Binance	6	5.733.637
Crypto4Pumps	11.716	-	Vip	Bittrex	47	491.395
Pump King Community	7.771	-	Vip	Bittrex	18	931.960
Crypto Family Pumps	4.449	5.299	Free	Cryptopia	28	23.800
Luxurious Pumps	6.020	-	Free	YoBit	16	4.997
AltTheWay	7.333	-	Free	YoBit	89	700

Tabella 1 - Analisi dei gruppi di Pump and Dump trovati su Telegram e Discord

Nel paper di Li et al. [2] si trovano alcune evidenze di come sono organizzati questi gruppi. Dal loro studio emerge l'importanza di *Telegram* come veloce strumento di comunicazione; infatti, essendo un'app di messaggistica permette di mettersi in contatto con gli utenti iscritto al gruppo con estrema rapidità e facilità. Le comunicazioni da parte degli organizzatori stabiliscono data, ora e per ultima viene svelata la criptovaluta coinvolta e vengono inviati numerosi promemoria per ricordare agli utenti di non perdere l'evento.

Il fatto di svelare il nome della criptovaluta "bersaglio" all'ultimo momento permette da un lato agli organizzatori di effettuare le operazioni di acquisto prima che vengano acquistate dagli altri utenti e di non fornire sufficiente tempo ad eventuali non iscritti al canale di partecipare all'evento, poiché questo dura pochissimi minuti.

Anche nello studio [2], prima menzionato, si trova evidenza del fatto che questi gruppi hanno una struttura gerarchica ben definita: infatti è possibile acquistare una sottoscrizione premium per essere avvertiti prima degli altri di quale sarà la criptovaluta bersaglio.

Spesso si riesce a guadagnare un ruolo importante nella gerarchia del gruppo anche portando altre persone ad iscriversi e quindi aiutando a diffondere l'evento. Infatti, la riuscita di un evento di *Pump and Dump* è data soprattutto dal numero di partecipanti: più sono e più si riesce ad alterare il prezzo della criptovaluta.

Inoltre, con un alto numero di partecipanti è sufficiente che questi investano poco affinché il prezzo della criptovaluta vari, cosa che è più complicata quando pochi utenti acquistano un volume basso di titoli.

Per cercare di aumentare il numero di partecipanti spesso questi eventi vengono pubblicizzati su più gruppi contemporaneamente o su più social contemporaneamente. Si trova infatti evidenza sia del fatto che alcuni gruppi operano contemporaneamente sia su *Telegram* che su *Discord* e anche del fatto che gli stessi eventi vengono pubblicizzati su più canali contemporaneamente.

Quando si accede a questi gruppi si ha la visibilità sull'intera conversazione presente nel gruppo e in qualche modo permette a questi canali di "accreditarsi".

Telegram permette anche di vedere da quanti utenti un determinato messaggio è stato visualizzato, cosa che sicuramente aiuta gli organizzatori a coordinare la comunicazione, un esempio ripreso dal paper [2] è riportato nella Figura 4.

Per rendere l'idea di come funziona concretamente il fenomeno è stato estratto un caso di *Pump and Dump* ripreso dal dataset utilizzato per l'analisi fatta in questa tesi.



Figura 3 – Andamento del prezzo della criptovaluta AppCoins e il numero di unità scambiate, prima, dopo e durante l'evento

Nella Figura 3 è riportato l'andamento del prezzo della criptovaluta Appcoins (APPC) e dell'ammontare acquistato; l'evento in questione è stato pianificato per il 4 Novembre 2020 alle ore 17.

Prima dell'inizio dell'evento, alle 16:58:26 il prezzo della criptovaluta si attesta a 0,00000249 BTC corrispondenti a 0,09 \$, anche nei giorni precedenti al giorno dell'evento non si notano particolari fluttuazioni di prezzo, alle 17:00:04 il prezzo

della criptovaluta passa in 18 secondi a 0,11 \$ generando un aumento di circa il 25%. Il volume di criptovaluta scambiata riportato da Binance passa dalle 581 unità delle 16:58:20 a 3,71 milioni di unità nei primi 25 secondi intercorsi tra le 17:00:00 e le 17:00:25.

Nella successiva ora il volume della criptovaluta scambiata subisce delle oscillazioni, raggiungendo le 2,5 mln di unità scambiate in due momenti: alle 17:15:50 e alle 17:31:40. Come il volume scambiato anche il prezzo nell'ora successiva all'evento subisce delle importanti fluttuazioni passando dai 0,11 \$ raggiunti all'inizio dell'evento per poi tornare al livello iniziale di 0,09 \$ alle 17:23:50, successivamente il prezzo scende ulteriormente fino al valore di 0,07 \$ perdendo più del 30% del suo valore e addirittura sotto al livello di prezzo iniziale in meno di un'ora.

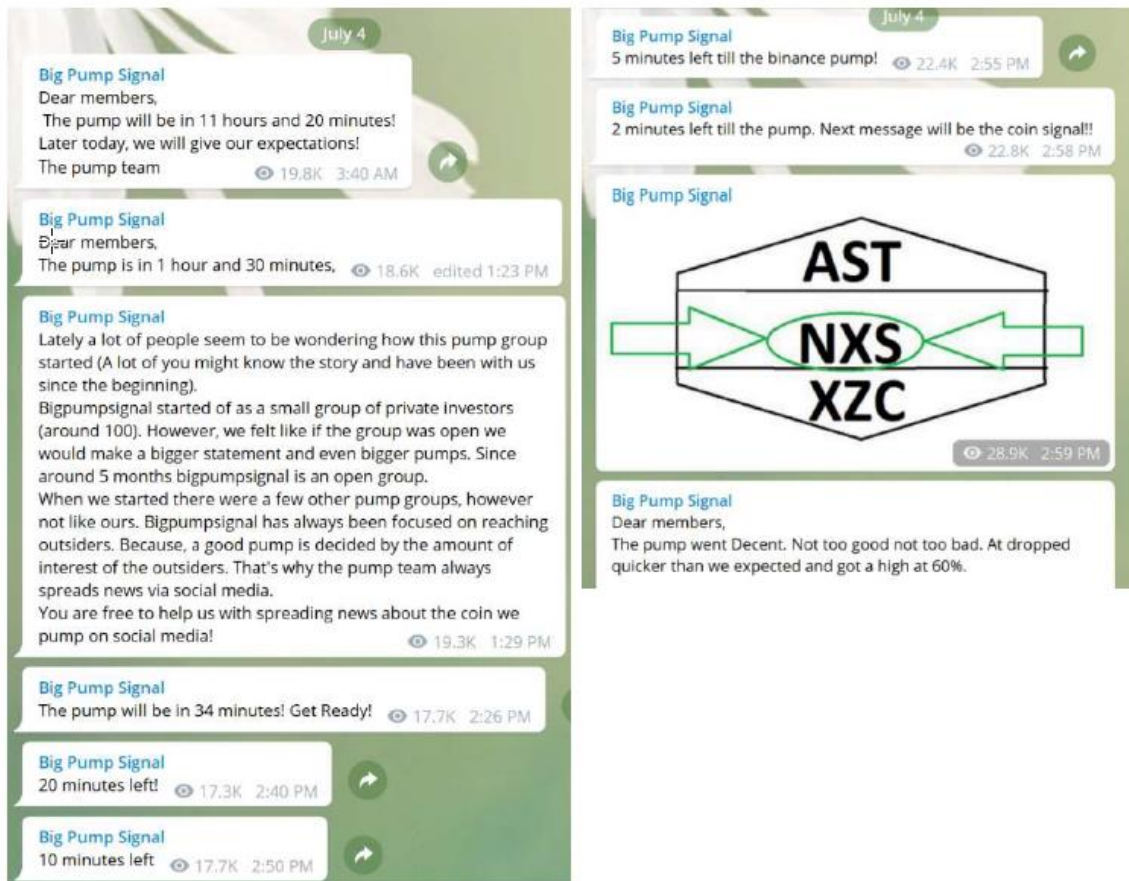


Figura 4 – Frammento di conversazione ripresa dal gruppo Telegram in cui viene annunciato un evento Pump and Dump

Lavori correlati al fenomeno

Come anticipato anche nel precedente paragrafo il *Pump and Dump* è un fenomeno tutt'altro che nuovo, che ricorre anche nel mercato azionario tradizionale.

Esistono molti studi che ne cercano di cogliere le peculiarità e gli aspetti affini, uno di questi è condotto da F. Allen et al. [3] i quali identificano tre categorie di schemi di manipolazione del mercato:

- i. uno basato sulle azioni che vengono fatte per manipolare il valore, almeno apparentemente, di un titolo;
- ii. uno basato sulla diffusione di informazioni false o fuorvianti;
- iii. uno basato sulla semplice compravendita di un titolo azionario.

Il *Pump and Dump* è il risultato di una combinazione del secondo e del terzo tipo di manipolazione del mercato.

Mei et al. [4] dimostrano invece come sia possibile mettere in piedi uno schema *Pump and Dump* solo sfruttando le asimmetrie informative degli investitori e confutano la loro tesi prendendo i casi del fenomeno segnalati dall'organo di vigilanza sugli scambi azionari statunitense (SEC) tra il 1980 e il 2002.

Mentre invece i ricercatori Khwaja et al. [5] dimostrano come i mercati emergenti siano più soggetti agli eventi di *Pump and Dump*, perché spesso adottano regolazioni meno stringenti; anche il nostro caso di studio rientra tra questi.

In uno studio condotto dall'Università di Innsbruck [6] invece, si osserva come il successo di Internet nei primi anni 2000 abbia portato alla nascita di una miriade di nuovi schemi attuati via mail e non più via telefono come in passato. La possibilità di inviare milioni di mail con un clic permette di raggiungere un maggior numero di utenti e quindi di alterare il prezzo di un titolo molto più facilmente rispetto al passato.

Come dimostrato da uno studio di Siering del 2013 [7] gli schemi realizzati tramite l'invio di mail ai consumatori finali riscuotono ancora oggi un discreto successo nonostante i provvedimenti presi dalle autorità di regolamentazione.

Negli ultimi anni sono stati pubblicate delle ricerche anche sugli eventi *Pump and Dump* nel mercato delle criptovalute, Gandal et al. [8] dimostrano come la prima "impennata" del prezzo del Bitcoin è stata generata da una manipolazione del mercato. Analizzando il noto dataset della piattaforma Mt.Gox rilevano infatti delle attività sospette condotte da due individui: "Markus bot" e "Willy bot". L'obiettivo di questi due soggetti era di comprare Bitcoin per aumentarne artificialmente il valore e il volume giornaliero.

Krafft et al. [9] indagano sulle abitudini comportamentali degli utenti della piattaforma Cryptsy. Nel loro paper dimostrano che anche l'acquisto di piccoli volumi di criptovaluta possono influenzare il mercato. Per acquistarle usano dei bot e ne concludono che i trader sono maggiormente interessati alle criptovalute sulle quali sono state registrate delle attività più recentemente.

Li et al. [2] conducono un'analisi empirica dei dati estratti da eventi di *Pump and Dump* avvenuti su Binance, Bittrex e Yobit, focalizzandosi sugli aspetti economici. I ricercatori mostrano come il *Pump and Dump* porta ad aumenti del prezzo e dei volumi nel breve termine, oltre che ad un'elevata volatilità, seguite da un'inversione della tendenza nei pochi minuti successivi all'evento. Dimostrano inoltre che il guadagno degli investitori è molto influenzato da quanto velocemente viene svelata loro la criptovaluta coinvolta nell'evento, per questo gli investitori esterni si trovano ad essere più svantaggiati rispetto agli altri.

Molto importante è il contributo di Kamps et al. [1] che si avvicinano per primi alla rilevazione dei fenomeni di *Pump and Dump*, questi studiosi descrivono il fenomeno e sottolineano la crescita dei mercati delle criptovalute, segnalando che all'epoca del loro paper (2018) corrisponde a 300 miliardi di \$, paragonabile al PIL della Danimarca. Inoltre, lanciano un monito sui nascenti fenomeni di *Pump and Dump* e sull'esigenza di trovare una regolamentazione per questi strumenti finanziari. Proprio perché uno tra i primi studi in questo ambito, i ricercatori sottolineano quindi la possibilità di validare il loro modello proprio per la difficoltà nel reperire i dati.

Ultimo, ma non per importanza, è il contributo che danno La Morgia et al. nel loro paper [10] che cercano di individuare i fenomeni *Pump and Dump* in tempo reale utilizzando tecniche come il Random Forest con i quali ottengono delle ottime performance di previsione degli eventi di *Pump and Dump*. Un altro contributo

molto importante da parte di questi ricercatori è la costruzione di un dataset per permettere un approfondimento su questo tema, come segnalato infatti da Kamps et al. [1], uno dei principali ostacoli allo studio del *Pump and Dump* nel mercato delle criptovalute è proprio la mancanza di un dataset utile a condurre un'analisi di questo tipo.

Rilevazione ed analisi degli eventi Pump and Dump

Panoramica del dataset

Il dataset utilizzato per l'analisi si compone di due parti:

- un file raccoglie tutti gli eventi di *Pump and Dump* estratti tramite tecniche di *Text Mining* da gruppi *Telegram* scelti per individuare gli eventi di *Pump and Dump*. Questo file contiene le criptovalute protagoniste dell'evento, il gruppo dal quale è stato individuato l'evento, data e ora nella quale dell'evento e piattaforma sulla quale l'evento verrà realizzato;
- l'altra parte del dataset invece è composta da un file, per ciascun evento, che contiene le transazioni di acquisto e di vendita della criptovaluta una settimana prima dell'evento e una settimana dopo. Ciascun file contiene la criptovaluta interessata, il timestamp e la data della transazione in formato data e ora, il tipo operazione (acquisto o vendita), il prezzo in Bitcoin della criptovaluta al momento della vendita o dell'acquisto, il numero di unità di criptovaluta scambiate e il valore della criptovaluta scambiata in Bitcoin (BTC).

Eventi di Pump and Dump

Le informazioni sugli eventi e sui gruppi sono state ricavate prima effettuando una ricerca utilizzando parole chiave come “Pump”, “Dump” e “Signal” su *Telegram*,

*Twitter, Reddit, BitcoinTalk*¹⁶ e *PADL*¹⁷. Una volta individuati i gruppi sono state estrapolate e analizzate le conversazioni di 20 gruppi *Telegram* in un periodo compreso tra luglio 2017 e gennaio 2021. Dall'analisi delle conversazioni su *Telegram* sono stati estratti in totale 904 eventi di *Pump and Dump*.

Gli eventi individuati si sono verificati su 5 piattaforme diverse: Binance, Bittrex, Cryptopia, Kucoin e Yobit e vedono coinvolte 372 diverse criptovalute a bassa capitalizzazione.

Una volta ottenuto il dataset si è svolta un'analisi esplorativa per cercare di scoprire alcuni *pattern* comuni ai *Pump and Dump*, questo da un lato è utile a capire quali siano le strategie che i gruppi mettono in atto e dall'altro potrebbero essere uno spunto per futuri approfondimenti.

La prima cosa che salta all'occhio è che data la numerosità degli eventi individuati (1110), che è maggiore del numero di eventi distinti, è facile intuire che molti di questi vengano promossi su più gruppi contemporaneamente. Questo permette agli organizzatori di raggiungere un numero di partecipanti maggiore e quindi generare un guadagno maggiore. Ad avvalorare questa ipotesi sono le Figure 5 e 6, infatti se raggruppiamo il numero di *Pump and Dump* per nome della

¹⁶ BitcoinTalk

<https://bitcointalk.org/>

¹⁷ PADL

<https://padl.mine.nu.siteindices.com/>

criptovaluta, data e ora su Binance, la piattaforma sulla quale si nota il divario maggiore, il numero di eventi passa da 520 a 338.

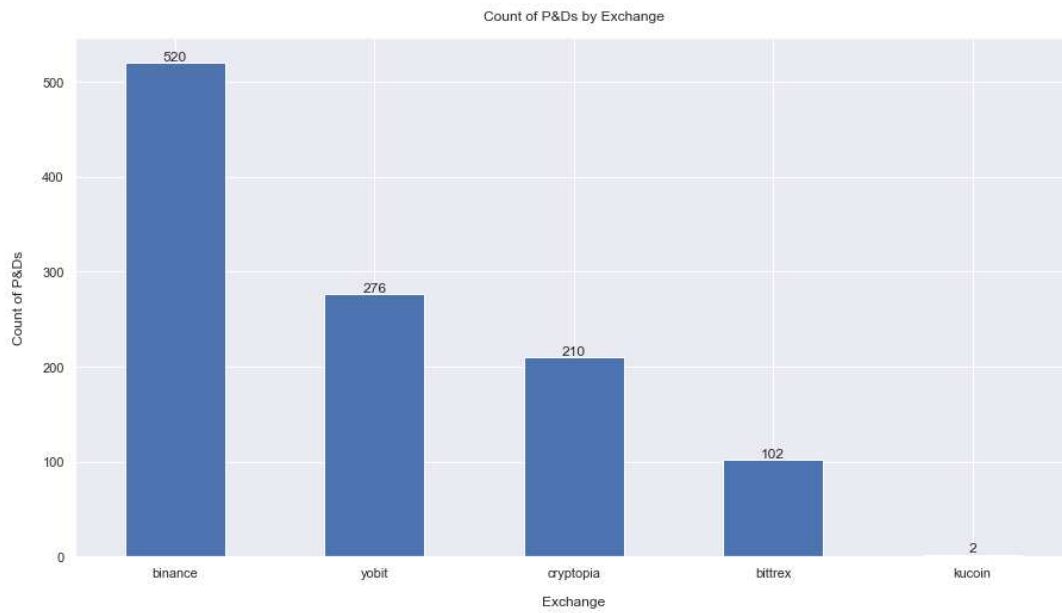


Figura 5 – Analisi dei gruppi di Pump and Dump trovati su Telegram e Discord

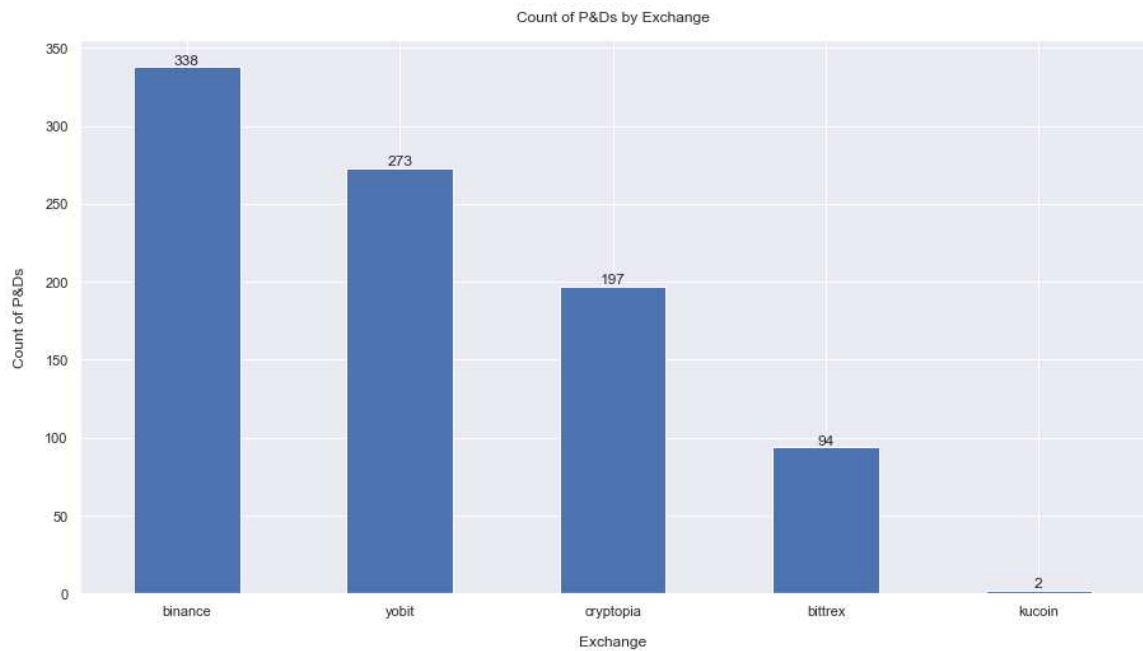


Figura 6 – Numero di eventi Pump and Dump individuati omettendo quelli promossi da più gruppi

Questo pattern è riscontrabile sulla criptovaluta CureCoin (CURE) sulla quale vengono promossi tre eventi distinti:

- Il primo pianificato per il 3 ottobre 2017 e promosso solamente dal gruppo Crypto4Pumps;
- Il secondo invece, pianificato per il 19 marzo 2019 è promosso contemporaneamente dai gruppi Cryptopia pump squad e Big Pump Family;
- Il terzo, infine, pianificato per il 12 dicembre 2019 e promosso da Cryptopia Pump Squad, Trading Crypto Guide e Big Pump Family.

In modo analogo viene fatto su altre criptovalute come Cortex (CXTC), Bored Coin (BRD) ed altre ancora.

Da questo primo esempio è possibile notare come solitamente i gruppi che organizzano un evento tendono a prendere di mira la stessa criptovaluta più volte prima di passare ad un'altra.

Nel periodo preso in analisi è possibile anche riscontrare che le stesse criptovalute vengono prese di mira più volte dai gruppi che organizzano gli eventi di *Pump and Dump*. Analizzando la criptovaluta Ambire AdEx (ADX) è possibile notare che è soggetta a tre eventi di *Pump and Dump*: una prima volta il 5 luglio 2019 alle ore 17 poi da un altro gruppo il 26 luglio 2019 alle ore 17:59 e infine da tre gruppi diversi circa un anno dopo, il 7 luglio 2020 alle ore 17. Non è stato possibile dimostrarlo con certezza ma è plausibile che i tre gruppi promotori dell'evento del 2020 siano legati in qualche modo o che gli amministratori siano addirittura gli stessi, poiché come ribadito più volte, il successo di un evento di *Pump and Dump* è dettato anche dal numero di partecipanti che ne prendono parte. In tutti e tre i casi la piattaforma sulla quale la criptovaluta viene scambiata è Binance questo perché spesso le piattaforme impongono dei costi per il trasferimento del proprio portafoglio di crypto ad un'altra piattaforma e pertanto sarebbe poco conveniente per gli organizzatori.

Analizzando la criptovaluta Aurora (AOA) invece, possiamo notare che è soggetta a tre eventi diversi ma organizzati dallo stesso gruppo.

Per dare evidenza di questo “modus operandi” dei gruppi presi in analisi nella Tabella 2 sono stati riportati il numero di eventi *Pump and Dump* organizzati da ciascun gruppo e il numero di diverse criptovalute scelte dagli stessi. Se prendiamo il primo gruppo “AltTheWay” (ATW) notiamo che a fronte di 253 eventi promossi sono soltanto 122 le criptovalute coinvolte. Questo *pattern* è riscontrabile anche nel secondo e nel terzo gruppo (BPF e CPI).

Gruppo	# Eventi organizzati	# Criptovalute distinte	Exchange principale
ATW	253	122	Yobit
BPF	230	137	Binance
CPI	181	112	Cryptopia,Binance
CPS	58	48	Cryptopia,Binance
TCG	46	35	Binance
C4P	45	34	Bittrex
BPS	41	33	Binance
FCS	36	33	Binance
CW	33	27	Binance
MPG	32	24	Binance
BPG	23	10	Binance
TWP	23	17	Binance
TCC	22	21	Yobit
LUX	17	17	Yobit
WCG	17	17	Binance
CCS	14	12	Binance
PKG	14	13	Bittrex
CCB	12	12	Binance
CP	8	5	Binance
SE	5	5	Binance

Tabella 2 - Numero di eventi Pump and Dump organizzati da ciascun gruppo e numero di criptovalute

Delle cinque piattaforme di Exchange citate prenderemo in analisi gli eventi di *Pump and Dump* effettuati su Binance, questo perché Cryptopia ha presentato istanza di fallimento nel 2019 pertanto non fornisce più dati sulle transazioni da poter estrarre e inoltre perché la maggior parte dei gruppi presi in analisi operano prevalentemente su Binance.

Transazioni

Come anticipato nel precedente paragrafo il dataset delle transazioni è composto da una serie di file contenenti tutti le stesse variabili:

- Prezzo in Bitcoin (BTC);
- Unità di criptovaluta acquistate o vendute;
- Valore della transazione in Bitcoin (BTC).

Pertanto, l'analisi verrà eseguita su un subset di transazioni riferite a diversi eventi *Pump and Dump*.

Per analizzare l'evoluzione del fenomeno e le possibili implicazioni che nel tempo possono avere su un progetto crypto è stata scelta la criptovaluta Bored Coin (BRD) coinvolta in 14 eventi distinti tra il 2018 e il 2020.

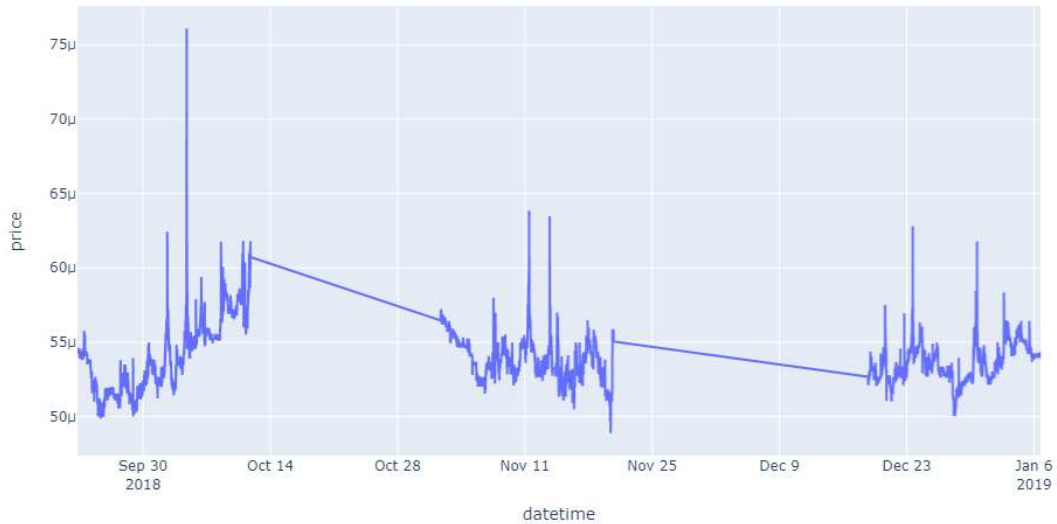


Figura 7 – Variazioni di prezzo della Bored Coin durante i Pump and Dump del 2018

Analizzando gli eventi avvenuti il primo anno, mostrati nella Figura 7, possiamo notare come rispetto al livello iniziale il prezzo subisce un calo dopo il primo evento, in corrispondenza del primo evento di *Pump and Dump* il prezzo di un'unità di criptovaluta passa da un minimo di 208,07 \$ il 25 settembre 2018 a 317,70 \$ il 4 ottobre 2018, cioè nel giorno dell'evento per stabilizzarsi a 257 \$ il giorno 11 ottobre 2018 (tenendo conto del valore del *Bitcoin* nel periodo di riferimento) e vengono acquistate 877771 unità di criptovaluta nei primi 5 minuti dell'evento per un valore di scambio di circa 280.000 \$. Analizzando lo stesso arco temporale sulle operazioni di vendita invece riscontriamo che sono state scambiate 550254 unità per un valore totale di 181.106 \$. L'incremento nei

volumi di scambio è chiaramente visibile nella Figura 8 che riporta le unità acquistate prima, durante e dopo il primo evento di *Pump and Dump*.

Nel successivo evento, che ricorre il 13 novembre 2018 è possibile notare il prezzo della criptovaluta parte da un livello più basso, circa 242 \$ e raggiunge un valore massimo di 271 \$ durante il tempo di monitoraggio.

Analizzando lo scambio riscontriamo che sono state acquistate 731.465 unità di criptovaluta acquistata, nei primi cinque minuti dell'evento, per un valore di circa 231.000 \$. Sul lato delle vendite si riscontra una flessione rispetto al primo evento con 422.033 di unità vendute per un valore di 133.000 \$. Il calo del valore dell'operazione è dovuto anche alla flessione delle criptovalute nel periodo di riferimento, risale al 2018 a delle indiscrezioni rilasciate dal network americano CBS, che avvertiva di una frode nel mercato delle criptovalute.¹⁸

Questo fattore fa emergere in modo molto forte l'instabilità del valore delle criptovalute nel sostenere le fughe di notizie.

¹⁸ 2018 Cryptocurrency bubble

https://en.wikipedia.org/wiki/Cryptocurrency_bubble#:~:text=By%20September%202018%2C%20cryptocurrencies%20collapsed,value%20in%20the%20previous%20week.

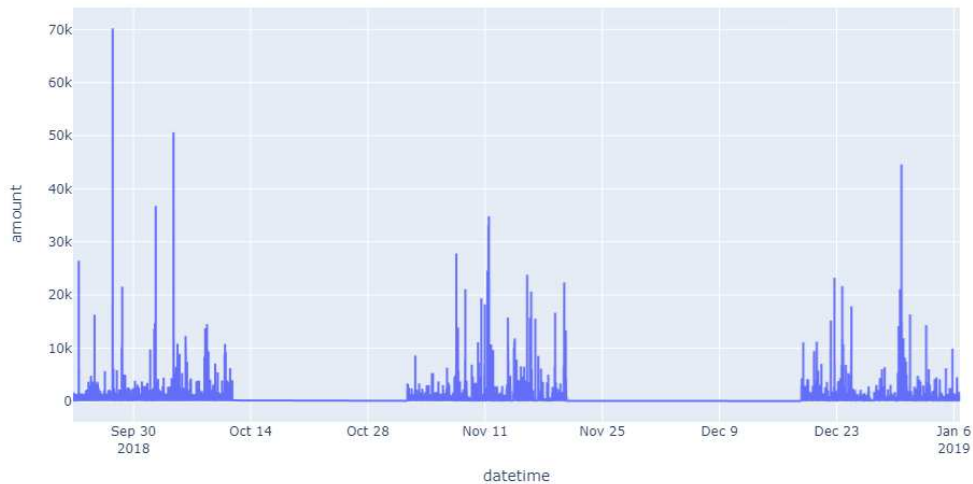


Figura 8 – Numero di unità scambiate durante i Pump and Dump del 2018

Il record negativo viene raggiunto nel terzo e ultimo evento del 2018, quando il *Bitcoin* raggiunge il valore unitario di 3739,31 \$, il numero di unità scambiate non si discosta molto da quello dell'evento precedente (657.861 unità acquistate e 460.567 vendite) tuttavia il valore delle unità acquistate si attesta intorno ai 120.684 \$ e quello di quelle vendute ad 81.548 \$.

Nel 2019 la *Bored Coin* è stata bersaglio di cinque eventi *Pump and Dump*, il primo risale al 17 marzo, nel periodo prima e dopo l'evento non si registrano particolari fluttuazioni: facendo un'analisi più profonda si sono rilevati acquisti per 370.177 unità a fronte di un valore di 76.245 \$ e vendite per 161.120 unità per un valore di 33.102 \$. È possibile, tuttavia, notare un vertiginoso aumento del prezzo della moneta il 12 marzo, questo picco potrebbe essere stato causato da un

altro evento *Pump and Dump*, tuttavia, nel dataset utilizzato per l'analisi non vi sono elementi che lo possano dimostrare con certezza poiché il gruppo o i gruppi che l'hanno realizzato non sono presenti nel nostro dataset.

Sull'onda del generale aumento del prezzo delle criptovalute¹⁹ (testimoniato anche dal grafico ripreso da Google Finance e riportato nella Figura 10) durante il successivo periodo dell'anno nonostante il prezzo della criptovaluta, gli eventi *Pump and Dump* si sono mostrati molto meno efficaci, effettuando un confronto sulle variazioni di prezzo tra le Figure 7 e 9 che riportano gli andamenti del prezzo della criptovaluta negli eventi del 2018 e 2019, le stesse differenze si evidenziano confrontando le Figure 8 e 11 che riportano le unità di criptovaluta acquistate. È molto probabile che molti eventi in successione generino sfiducia nella community se gli utenti non riescono ad ottenere i guadagni sperati.

¹⁹ Andamento del mercato delle criptovalute nel 2019

<https://it.cointelegraph.com/news/someone-say-bear-market-bitcoin-sets-3-huge-new-records-in-2019>



Figura 9 – Fluttazioni del prezzo della Bored Coin durante i Pump and Dump del 2019

Quest'evidenza è confermata anche da La Morgia et al. [10] che analizzando lo storico delle conversazioni delle community di *Pump and Dump* riscontrano insoddisfazione nei commenti di alcuni membri. Osservando il grafico del valore delle transazioni in Bitcoin degli eventi realizzati nel 2019 (Figura 12), riscontriamo un calo rispetto a quelli verificatisi nel 2018. Analizzando i primi cinque minuti dell'evento del 25 maggio 2019, come fatto nei precedenti eventi, si riscontra un volume di soltanto 40.592 unità di *Bored Coin* acquistate a fronte di un valore di 16.000 \$. Questo scarso successo dell'evento potrebbe essere legato sia al numero di partecipanti al gruppo organizzatore (che in questo caso è *Alt The Way*) e sia al fatto che questo gruppo è più attivo sulla piattaforma di scambio Yobit rispetto a quella utilizzata in questo caso (Binance), come mostrato nella Tabella 1, riportata precedentemente.

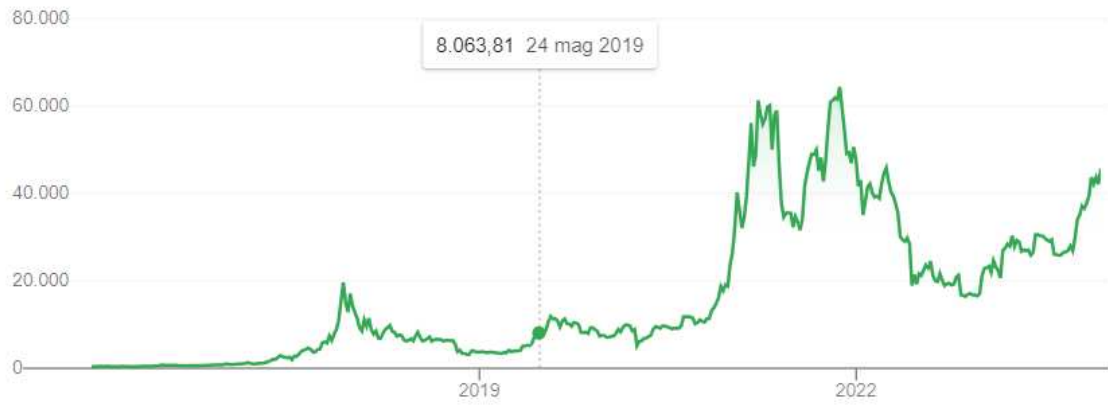


Figura 10 – Andamento storico del prezzo del Bitcoin nel corso degli anni (2015 – 2024)

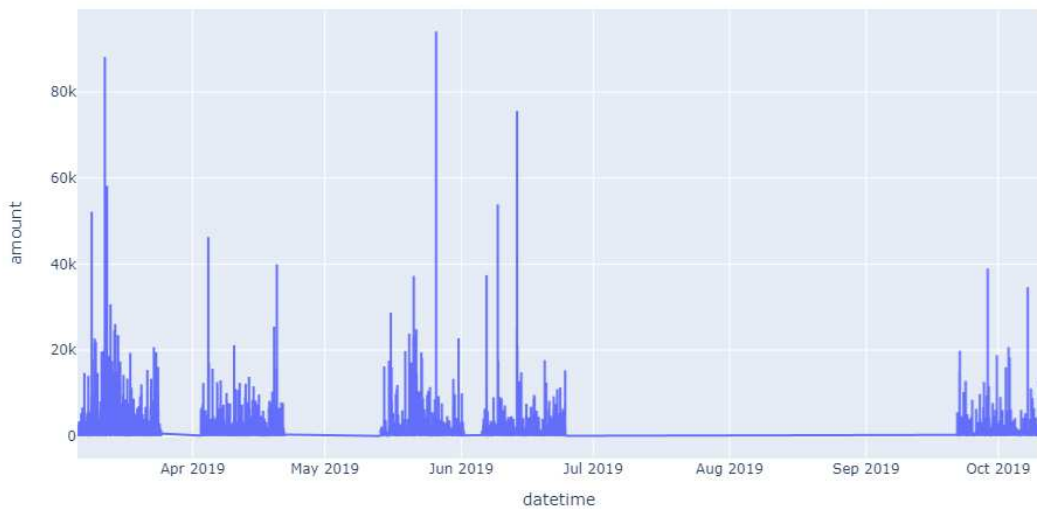


Figura 11 – Numero di unità scambiate durante i Pump and Dump del 2019

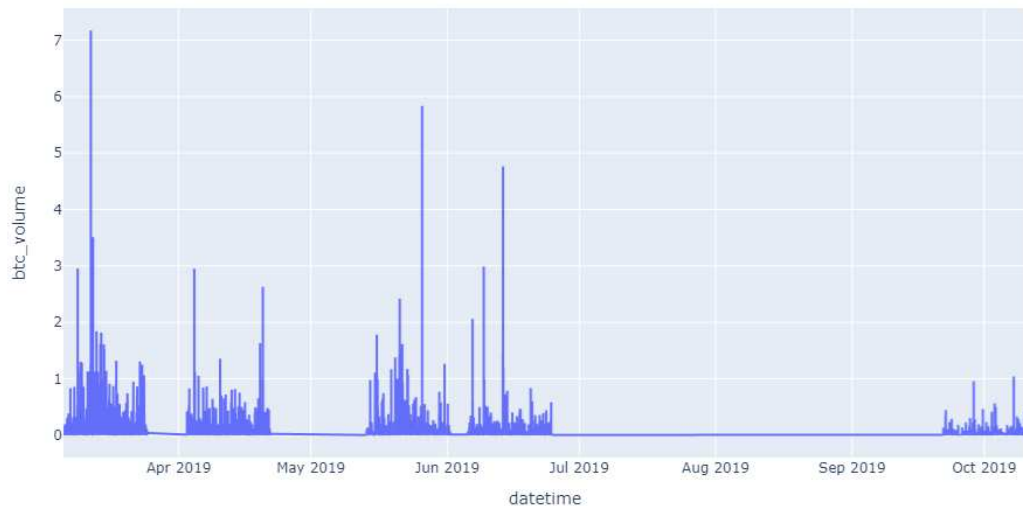


Figura 12 – Volumi di scambio in Bitcoin della Bored Coin nei Pump and Dump del 2019

Analizzando infine gli eventi del 2020 (riportato in Figura 13) è possibile notare come il prezzo sia notevolmente più basso rispetto a quello degli anni passati, questo è probabilmente dovuto all'incertezza economica causata dalla pandemia di COVID-19 del 2020. La correlazione positiva tra la situazione sanitaria ed economica globale e l'andamento del mercato delle criptovalute è dimostrata anche da Sarkodie et al. nel loro paper [11]. I ricercatori affermano che le criptovalute subiscono le conseguenze negative del mercato azionario tradizionale in caso di eventi catastrofici come una pandemia; tuttavia, possono essere utilizzati dagli investitori per diversificare il loro portafoglio azionario e difendersi dall'inflazione poiché le criptovalute come il Bitcoin hanno una

quantità fissa mentre la quantità di denaro sul mercato è soggetto alle politiche delle banche centrali che può aumentarne o diminuirne la quantità.

Tornando al primo evento del 2020 si riscontra un acquisto di 490.967 per un valore di 94.274 \$ (considerando il prezzo medio del Bitcoin nella settimana dell'evento di 8033,07 \$) e vendite per un ammontare di 114,876 \$.

Negli eventi successivi il prezzo continua a scendere, raggiungendo dei picchi solo durante gli eventi di *Pump and Dump* e passando dal massimo di 182,62 \$ del primo evento ai 99,68 \$ dell'evento del 24 marzo 2020; negli eventi di novembre e dicembre i prezzi della criptovaluta raggiungono rispettivamente i massimi di 91,55 \$ e 54,45 \$ a segnare il minimo storico tra gli eventi analizzati e una diminuzione del 60% del suo valore da gennaio 2020.



Figura 13 – Fluttazioni del prezzo della Bored Coin durante i Pump and Dump del 2020

Se analizzando l'andamento del prezzo della criptovaluta Bored Coin si evidenzia un calo del prezzo rispetto agli eventi degli anni passati sul lato dell'ammontare acquistato il trend risulta di segno opposto, come mostrato anche in Figura 14, si riscontrano acquisti per 243.963 \$ nell'evento di novembre e 142.822 \$ in quello di dicembre

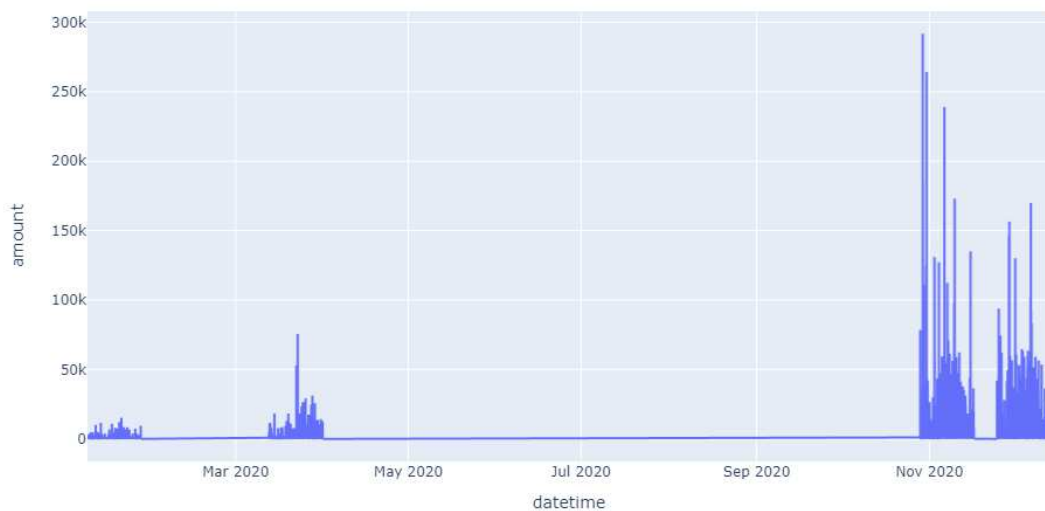


Figura 14 – Numero di unità scambiate durante i Pump and Dump del 2020

Definizione delle feature per la classificazione

Feature in letteratura

L'analisi parte dall'intuizione dell'importanza degli "ordini urgenti" negli eventi *Pump and Dump*. Questo tipo di acquisti vengono effettuati nel momento in cui viene svelata la criptovaluta bersaglio dell'evento, poco prima o poco dopo l'evento. La presenza di questo tipo di operazioni finanziarie è molto evidente nel primo evento che coinvolge la criptovaluta *Bored Coin* avvenuto il 4 ottobre 2018, analizzando la Figura 15 è possibile, infatti, riscontrare come in corrispondenza dell'ora prestabilita dell'evento l'andamento del prezzo subisca un considerevole aumento dato proprio dall'acquisto massivo del titolo dovuto al segnale di *Pump* da parte dei leader del gruppo.

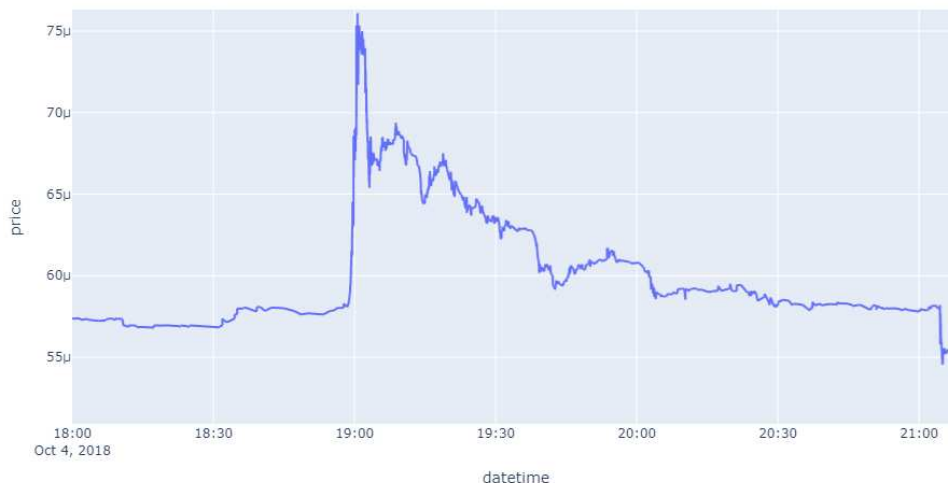


Figura 15 – Andamento del prezzo prima, durante e dopo l'evento

Di pari passo vanno i volumi di acquisto della criptovaluta riportati in Figura 16, se nell'ora prima dell'evento è possibile osservare un titolo poco mosso, subito dopo il segnale di Pump il numero di unità acquistate aumenta notevolmente così come il prezzo.

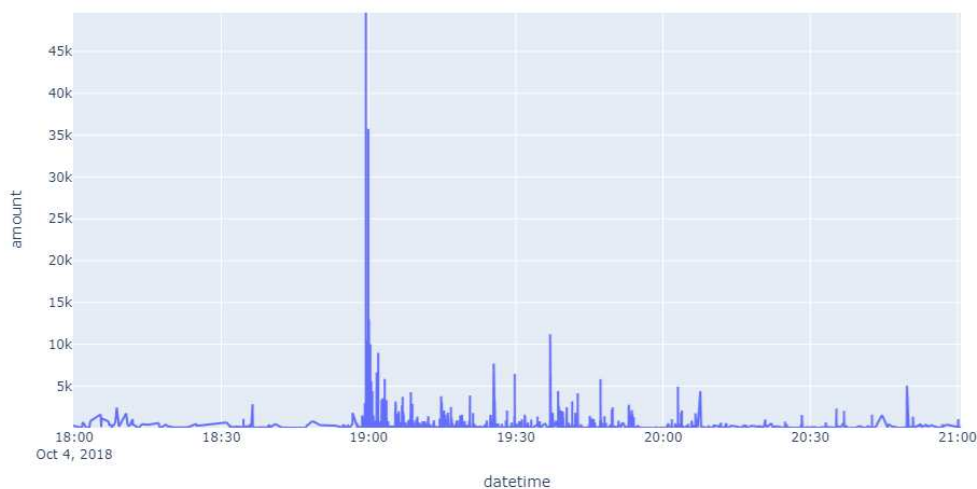


Figura 16 – Numero di unità scambiate nei momenti prima, durante e dopo l'evento

Per questo motivo, attraverso una serie di test, viene calcolata una finestra temporale che catturi le variazioni nel tempo dell'evento.

Le feature calcolate per l'analisi sono:

- **Standard Rush Orders:** deviazione standard mobile²⁰ del numero di ordini “urgenti” (rush) in ciascuna finestra temporale

²⁰Moving Standard Deviation formula

<https://www2.microstrategy.com/producthelp/current/FunctionsRef/Content/FuncRef/MovingStandardDeviationMovingStandardDeviation.htm>

$$x_k = \sqrt{\frac{\sum_{i=k-m+1}^k (x_i - \bar{x})^2}{m-1}} \quad \forall k = 1, \dots, n$$

- **Average Rush Orders:** media mobile²¹ del numero di ordini “urgenti” (rush) in ciascuna finestra temporale

$$SMA_k = \frac{1}{k} \sum_{i=n-k+1}^n x_i$$

- **Standard Trades:** deviazione standard mobile del numero di transazioni di acquisto e di vendita all’interno della finestra temporale

$$y_k = \sqrt{\frac{\sum_{i=k-m+1}^k (y_i - \bar{y})^2}{m-1}} \quad \forall k = 1, \dots, n$$

- **Standard Volumes:** deviazione standard mobile del valore delle transazioni in Bitcoin in ciascuna finestra temporale

$$z_k = \sqrt{\frac{\sum_{i=k-m+1}^k (z_i - \bar{z})^2}{m-1}} \quad \forall k = 1, \dots, n$$

- **Average Volumes:** media mobile del valore delle transazioni in Bitcoin in ciascuna finestra temporale

$$SMA_k = \frac{1}{k} \sum_{i=n-k+1}^n z_i$$

²¹ Simple Moving Average formula

https://en.wikipedia.org/wiki/Moving_average

- **StdPrice:** deviazione standard mobile del prezzo di chiusura della finestra temporale

$$p_k = \sqrt{\frac{\sum_{i=k-m+1}^k (p_i - \bar{p})^2}{m-1}} \quad \forall k = 1, \dots, n$$

- **AvgPriceMax:** media mobile del prezzo massimo in ciascuna finestra temporale

$$SMA_k = \frac{1}{k} \sum_{i=n-k+1}^n \max p_i$$

- **AvgPriceMin:** media mobile del prezzo minimo in ciascuna finestra temporale

$$SMA = \frac{1}{k} \sum_{i=n-k+1}^n \min p_i$$

- **Gt:** è una variabile che serve a classificare gli eventi *Pump and Dump*, assume valore 1 nel momento in cui si verifica l'evento e 0 in tutti gli altri casi

Il calcolo delle feature viene effettuato utilizzando diverse set di finestre temporali:

- Il primo set utilizza una finestra temporale da 25 secondi e una rolling frequency di 900;

- Il secondo utilizza una finestra temporale da 15 secondi e una rolling frequency di 900;
- Il terzo set riduce la finestra temporale a 5 secondi e una rolling frequency di 700.

Gli indicatori a media mobile sono stati utilizzati perché sono degli indici ampiamente utilizzati nelle analisi sugli andamenti dei mercati finanziari. Questo tipo di misure permettono di ridurre l'effetto delle fluttuazioni di prezzo di breve termine, inoltre, a seconda dei fenomeni che si vogliono studiare le ampiezze delle finestre temporali possono essere aumentate o ridotte. Generalmente se i valori sono sopra la media mobile il trend del titolo è al rialzo mentre se sono inferiori alla media mobile il trend è potenzialmente a ribasso.²²

Matrix Profile

Il Matrix Profile è una tecnica di analisi delle serie storiche molto utilizzata per problemi di Anomaly Detection. È stata introdotta per la prima volta nel 2016 dai ricercatori Eamonn Keogh della University of California Riverside e Abdullah Mueen della University of New Mexico, ha il vantaggio di essere molto efficiente, ha dei tempi di elaborazione non molto lunghi (a seconda dell'algoritmo che viene utilizzato) ed è indipendente dal dominio applicativo in cui viene utilizzato.

²²Utilizzo degli indici a media mobile nelle analisi di mercato

<https://www.investopedia.com/articles/active-trading/052014/how-use-moving-average-buy-stocks.asp>

Il Matrix Profile si compone di due elementi:

- 1) Un vettore di distanze euclidee standardizzate;
- 2) Un *profile index* che contiene l'indice della sequenza di distanze più vicina.

L'approccio *Matrix Profile* utilizza una finestra mobile di ampiezza "m". È possibile riassumere il funzionamento della tecnica in sette step di calcolo:

1) Definizione della finestra mobile:

È necessario scegliere la lunghezza della finestra mobile. Questa rappresenta la lunghezza delle sottosequenze che verranno confrontate tra loro. La scelta di questa lunghezza dipenderà dalla natura dei dati e dai pattern che si desidera individuare.

2) Calcolo delle sottosequenza mediana:

Per ciascuna finestra mobile, calcola la sottosequenza mediana, ovvero il valore medio della sotto-sequenza.

3) Calcolo delle distanze:

Calcola la distanza euclidea tra la sottosequenza corrente e tutte le altre sottosequenze nella finestra mobile. È possibile utilizzare altre metriche di distanza se necessario.

4) Calcolo del Profilo di Distanza:

A partire dai risultati delle distanze, si calcola il profilo di distanza. Questo profilo rappresenta la distanza minima per ciascuna finestra mobile rispetto a tutte le altre finestre mobili, come riportato nell'esempio in Figura 17.

Compare Subsequences

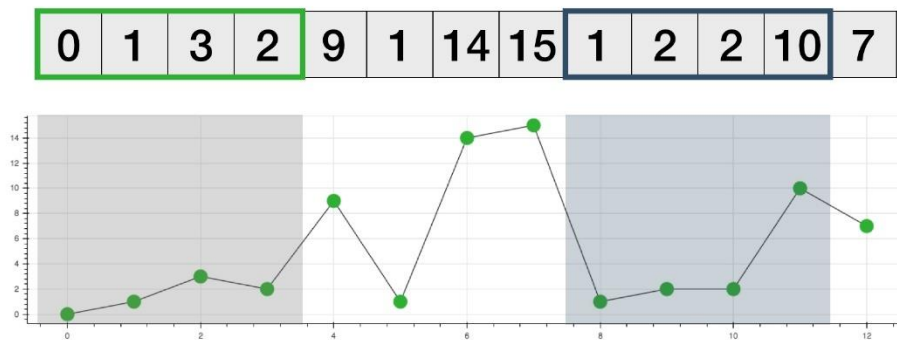


Figura 17 – Confronto delle distanze tra le sottosequenze di una serie storica di lunghezza $n=13$

5) Calcolo del Matrix Profile:

Per ottenere il *Matrix Profile* completo, si estende il profilo di distanza a coprire l'intera serie temporale. In pratica, si concatenano le distanze calcolate per ogni finestra mobile.

6) Identificazione dei Motif e delle Anomalie:

Individua i minimi locali nel *Matrix Profile*. Questi minimi indicano le sottosequenze più simili o i "motif" nella serie temporale. È possibile

utilizzare queste informazioni per identificare sia i pattern ricorrenti che le potenziali anomalie.

7) Setting della soglia di anomalie:

Trova i minimi locali nel *Matrix Profile*. Questi minimi indicano le sottosequenze più simili o i "motif" nella serie temporale. Si possono utilizzare queste informazioni per identificare sia i pattern ricorrenti che le potenziali anomalie.

Come mostrato nell'esempio in Figura 18, vengono detti "motif" i pattern che si ripetono mentre quelli che non si ripetono vengono detti "discord" o anomalie. Il *Matrix Profile* memorizza le distanze euclidee tra due o più sequenze temporali, maggiore è la distanza e maggiore è la probabilità che ci siano delle anomalie, se invece la distanza assume un valore molto basso (ad es. vicino allo zero) maggiore è la probabilità che sia simile ad un'altra sequenza temporale e quindi non sarà classificata come anomalia.²³

²³Introduzione al Matrix Profile

<https://towardsdatascience.com/introduction-to-matrix-profiles-5568f3375d90>

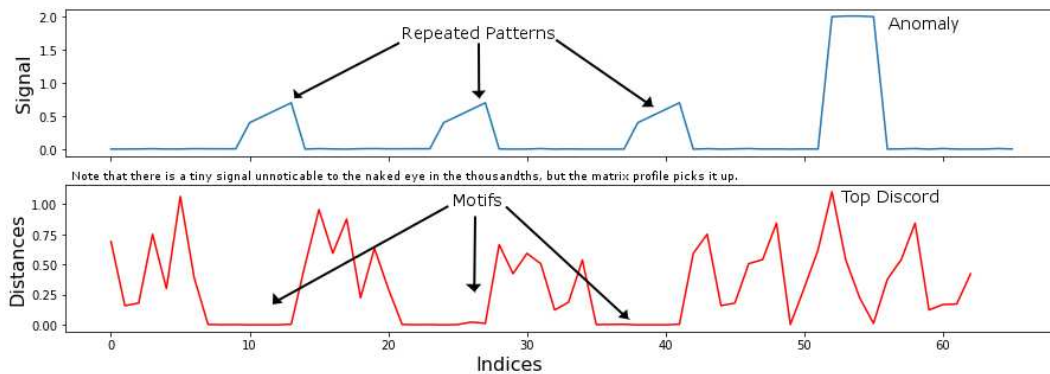


Figura 18 – Rilevazione dei motifs su un grafico delle distanze

Il calcolo del *Matrix Profile* può essere realizzato tramite diversi algoritmi, tra i più conosciuti ci sono:

- Naive: questo approccio è definito “brutale” ed è il meno efficiente di tutti per l’analisi di serie storiche molto ampie. Questo perché la complessità di calcolo cresce più che proporzionalmente all’aumentare delle osservazioni.
- STAMP: è uno dei primi algoritmi implementati dal gruppo di ricerca di Keogh. Lo STAMP è un algoritmo “anytime”²⁴ è in grado cioè di restituire una soluzione valida anche se viene interrotto anticipatamente. Questo algoritmo in particolare viene utilizzato con le serie storiche univariate.
- STOMP: è un’evoluzione dell’algoritmo STAMP altamente efficiente e facilmente scalabile, si basa su una tecnica di suddivisione ad albero (BOP

²⁴Definizione di Algoritmo Anytime

https://it.wikipedia.org/wiki/Algoritmo_anytime

- Binary Order-preserving Partitioning) per ridurre la complessità computazionale del calcolo delle distanze.²⁵
- SCRIMP: combina l'efficienza e la velocità di calcolo dello STOMP con il fatto di essere un algoritmo "Anytime" dello STAMP.

Classificazione degli eventi di Pump and Dump

Una volta aver compreso il fenomeno e le tecniche utilizzate si è passati all'implementazione del classificatore degli eventi *Pump and Dump*.

Nella prima parte dell'esperimento sono state effettuate alcune prove utilizzando le feature di partenza di La Morgia et al. [10] in modo da avere una baseline con la quale confrontare le performance degli altri modelli implementati.

Le feature menzionate precedentemente state calcolate utilizzando delle finestre temporali di 5, 10 e 25 secondi mentre a differenza del paper è stato utilizzato, in aggiunta all'algoritmo *Random Forest*, l'*XGBoost* con 5 fold di cross validation. Per valutare la bontà dei modelli di classificazione sono state utilizzate: Precision, Recall ed F1-score.

<i>Classificatore</i>	<i>Finestra temporale</i>	<i>Recall</i>	<i>Precision</i>	<i>F1-score</i>
RF (5 folds)	25 s	88.9 %	96.6 %	92.6 %
RF (5 folds)	15 s	85.5 %	96.8 %	90.8 %

²⁵Algoritmo STOMP

https://matrixprofile.org/matrixprofiler/reference/mp_algos.html

RF (5 folds)	5 s	73.5 %	94.3 %	82.6 %
XGBoost (5 folds)	25 s	92.4 %	97.0 %	94.7 %
XGBoost (5 folds)	15 s	87.7 %	93.9 %	90.7 %
XGBoost (5 folds)	5 s	79.8 %	87.8 %	83.6 %

Tabella 3 – Performance dei due modelli di partenza

Dalle performance ottenute con il modello di partenza e riportate in Tabella 3, risulta che il modello più efficace nell'identificazione degli eventi risulta l'XGBoost. I risultati sono stati ottenuti prima calcolando le feature su tre diverse temporali e poi testando gli algoritmi Random Forest e XGBoost con 5 fold di Cross Validation per ottenere dei risultati più stabili.

Successivamente si è proceduto al calcolo delle distanze tra le osservazioni con la tecnica del *Matrix Profile* delle feature numeriche di partenza: prezzo, quantità acquistate o vendute e valore in *Bitcoin*.

Dal momento che il tempo di elaborazione per il dataset utilizzato per l'esperimento sono stati condotti prima alcuni test su un subset di eventi *Pump and Dump* per individuare il valore della finestra temporale ottimale e poi è stato applicato all'intero dataset.

Il subset utilizzato contiene 12 eventi *Pump and Dump* realizzati tra il 2019 e il 2020 su 4 diverse criptovalute: *Ambire AdEx (ADX)*, *Ambrosus (AMB)*, *AppCoin (APPC)* e *Ardor (ARDR)* mentre i risultati degli esperimenti effettuati su window size di diversa ampiezza, con i due diversi algoritmi di classificazione sono riportati in Tabella 4.

<i>Classificatore</i>	<i>Finestra temporale</i>	<i>Window size</i>	<i>Recall</i>	<i>Precision</i>	<i>F1-score</i>
RF (5 folds)	25 s	3	79.4 %	90 %	84.4 %
RF (5 folds)	15 s	3	77.7 %	92.1 %	84.3 %
RF (5 folds)	5 s	3	66.4 %	93.4 %	77.6 %
XGB (5 folds)	25 s	3	79.4 %	87.1 %	83.1 %
XGB (5 folds)	15 s	3	80 %	92.3 %	85.7 %
XGB (5 folds)	5 s	3	67.2 %	85.1 %	75.1 %
RF (5 folds)	25 s	10	79.4 %	93.1 %	85.7 %
RF (5 folds)	15 s	10	77.8 %	89.7 %	83.3 %
RF (5 folds)	5 s	10	63.3 %	93.1 %	75.3 %
XGB (5 folds)	25 s	10	76.5 %	89.6 %	82.5 %
XGB (5 folds)	15 s	10	77.8 %	92.1 %	84.3 %
XGB (5 folds)	5 s	10	67.2 %	88.6 %	76.4 %
RF (5 folds)	25 s	30	79.4 %	90 %	84.4 %
RF (5 folds)	15 s	30	75.5 %	97.1 %	85 %
RF (5 folds)	5 s	30	65.6 %	93.3 %	77.1 %
XGB (5 folds)	25 s	30	79.4 %	93.1 %	85.7 %
XGB (5 folds)	15 s	30	80 %	92.3 %	85.7 %
XGB (5 folds)	5 s	30	67.8 %	87 %	85.7 %
RF (5 folds)	25 s	50	79.4 %	90 %	84.4 %
RF (5 folds)	15 s	50	75.5 %	91.9 %	82.9 %
RF (5 folds)	5 s	50	64.1 %	93.2 %	75.9%
XGB (5 folds)	25 s	50	79.4 %	96.4 %	87.1 %
XGB (5 folds)	15 s	50	80 %	94.7 %	86.7 %
XGB (5 folds)	5 s	50	67.2 %	87.7 %	76.1 %
RF (5 folds)	25 s	100	79.4 %	90 %	84.4 %
RF (5 folds)	15 s	100	77.8 %	94.6 %	85.4 %
RF (5 folds)	5 s	100	64.8 %	93.2 %	76.5 %
XGB (5 folds)	25 s	100	73.5 %	89.3 %	80.6 %
XGB (5 folds)	15 s	100	80 %	94.7 %	86.7 %
XGB (5 folds)	5 s	100	68.7 %	88 %	77.2 %

Tabella 4 – Performance degli esperimenti condotti per individuare la corretta window size

Il secondo obiettivo di questo test era quello di cercare di capire l'impatto della window size. Come riportato anche da Imani et al. [12] la window size è un parametro fondamentale nell'analisi delle serie storiche e può avere degli impatti negativi se non scelta in modo corretto o non si ha sufficiente esperienza sul dominio dei dati dell'analisi. Tenendo ferma la considerazione che il *Pump and Dump* è un evento che si verifica in modo molto veloce (come si è visto il prezzo di una criptovaluta può passare dal massimo al minimo storico in un'ora o meno) si è ipotizzato di non impostare una window size eccessivamente grande, questo anche per cercare di tenere a bada possibili altri eventi organizzati da gruppi non identificati.

Considerando quanto detto sopra e facendo un trade-off tra i migliori risultati ottenuti dalle tre misure, cercando di non sfavorire nessuna delle tre si è optato per un calcolo del *Matrix Profile* con una *window size* di ampiezza 30 ottenendo le performance riportate in Tabella 5.

<i>Classificatore</i>	<i>Finestra temporale</i>	<i>Window size</i>	<i>Recall</i>	<i>Precision</i>	<i>F1-score</i>
RF (5 folds)	25 s	30	68 %	84.6 %	75,5 %
RF (5 folds)	15 s	30	63 %	83.7 %	71.9 %
RF (5 folds)	5 s	30	50.6 %	85.5 %	63.6 %
XGB (5 folds)	25 s	30	67.8 %	79.3 %	73.1 %
XGB (5 folds)	15 s	30	70.3 %	79.7 %	74.7 %
XGB (5 folds)	5 s	30	76.8 %	82.3 %	79.5 %

Tabella 5 – Performance ottenute con le feature calcolate con il Matrix Profile

Rispetto al modello di partenza le performance sono inferiori, soprattutto utilizzando il Random Forest. È possibile riscontrare una diminuzione di 20 punti percentuali sulla Recall (sensibilità) ad indicare una diminuzione degli eventi classificati come *Pump and Dump* quando in realtà non lo sono e ad una diminuzione della Precision che identifica gli eventi classificati come eventi *Pump and Dump* rispetto a tutti gli eventi *Pump and Dump*. In particolare, la diminuzione di performance è maggiore quanto più diminuisce la finestra temporale presa in analisi. Per quanto riguarda l'XGBoost le performance sono inferiori a quelle calcolate utilizzando le feature di partenza ma risultano comunque migliori rispetto a quelle ottenute con il modello implementato con il Random Forest. Il miglior risultato tra i due algoritmi è quello ottenuto con XGBoost su una finestra temporale di ampiezza 5 secondi e riporta una Precision del 76.8%, una Recall dell'82.3% e un F1-score del 79.5%.

Discussione

Possiamo affermare che ci sia un aumento delle performance diminuendo l'ampiezza della finestra temporale utilizzando l'XGBoost, questo perché l'algoritmo risente molto degli outliers e del rumore che c'è nei dati. Come abbiamo visto anche nella panoramica del dataset è possibile che alcuni eventi non siano stati individuati in fase di preparazione dei dati poiché gli organizzatori non erano tra i gruppi finiti sotto la lente d'ingrandimento dell'analisi o semplicemente non avevano lasciato traccia di eventi *Pump and Dump* passati. Inoltre, l'algoritmo XGBoost, è anch'esso un modello basato su alberi decisionali ma a differenza del Random Forest in cui ciascun albero decisionale ha lo stesso peso e l'appartenenza ad una classe o all'altra è stabilito in base ad un numero di alberi che hanno lo stesso peso. Nel XGBoost gli alberi incidono sul risultato finale con pesi diversi e gli alberi successivi vengono addestrati per correggere gli errori del modello attuale. Il fatto che ciascun albero successivo corregga gli errori residui di quello precedente potrebbe spiegare anche la causa del miglioramento delle performance al diminuire della finestra temporale.

Il peggioramento delle performance del Matrix Profile potrebbe essere stato causato anche da una *window size* non sufficientemente ampia, ma poiché l'algoritmo si è rivelato computazionalmente molto laborioso, si è reputato più efficiente calcolare la miglior finestra mobile utilizzando un subset dei dati piuttosto che su una parte molto più sostanziosa di essi. Anche le irregolarità

riscontrate nei dati in fase di analisi, in alcuni casi si è infatti visto come i prezzi delle criptovalute subisse delle variazioni repentine senza essere prese come bersaglio da nessun gruppo. Le criptovalute hanno generalmente un andamento molto irregolare e sono spesso attacco di speculazione, anche quelle molto note come il Bitcoin e l'Ethereum, maggiormente sono esposte agli attacchi speculativi quelle a bassa capitalizzazione che sono soggette della nostra analisi.

Per quanto riguarda le limitazioni di questo approccio, c'è il fatto di non riuscire a comprendere fino in fondo alcune oscillazioni nei prezzi delle criptovalute e nelle unità scambiate; l'ipotesi è che potrebbero essere eventi organizzati da altri gruppi ma esistendo centinaia di gruppi che fanno *Pump and Dump* utilizzando piattaforme di comunicazione diverse è molto difficile individuarli tutti.

Un altro limite del modello realizzato è che esso è basato sulle variazioni repentine di prezzi o quantità, mentre se l'evento non riesce a raccogliere molti partecipanti è possibile che non venga individuato dal nostro classificatore.

Conclusione

In questa tesi si è fornita una panoramica sul funzionamento della tecnologia blockchain, che regola le operazioni di scambio tra le criptovalute, si è dato ampio spazio alle tecnologie coinvolte nella blockchain e si sono analizzate le sue componenti.

Successivamente si è introdotto il fenomeno del *Pump and Dump* spiegando le sue origini, si sono riportati alcuni casi emblematici nel mondo della finanza tradizionale e gli impatti che hanno avuto sui risparmiatori.

Si è data l'idea di quanto questo fenomeno sia distorsivo sull'andamento del mercato, sugli shock che può creare e si è cercato di comprendere l'importanza della tutela che la regolamentazione può dare ai comportamenti opportunistici sui mercati azionari.

Analizzando i gruppi presi in esame è stato possibile identificare gli eventi *Pump and Dump*, monitorando il numero di unità di criptovaluta acquistate e vendute e il prezzo è stato possibile capire quanto il fenomeno sia redditizio in alcuni casi per gli organizzatori e soprattutto quanto sia diffuso.

Infine, tramite il Matrix Profile si è cercato di individuare i pattern anomali nascosti nei dati e tramite il Random Forest e l'XGBoost si sono costruiti e confrontati diversi modelli di classificazione di questi eventi fraudolenti.

Il lavoro proposto in questa tesi potrebbe essere esteso in molte direzioni. Ad esempio, sarebbe interessante approfondire l'analisi introducendo degli algoritmi

di Computer Vision che cerchino di individuare le criptovalute bersaglio degli attacchi monitorando le conversazioni dei vari gruppi che li organizzano. Spesso, infatti, gli amministratori dei gruppi danno il segnale di Pump inviando un'immagine con il nome della criptovaluta per evitare i bot programmati per individuarlo prima degli altri. Questo potrebbe agevolare l'analisi delle conversazioni e potenzialmente portare a riscontrare più eventi di quelli si sono riusciti a trovare nell'analisi fatta.

Un altro ulteriore approccio potrebbe essere quello di studiare il legame che c'è tra gruppi diversi che promuovono lo stesso evento di *Pump and Dump* per capire se sono legate agli stessi membri e le implicazioni economiche che quest'unione scaturisce.

Infine, un ulteriore approccio potrebbe essere quello di capire quali esternalità negative hanno questi eventi su un progetto di criptovaluta.

Bibliografia

- [1] J. Kamps and B. Kleinberg, “To the moon: defining and detecting cryptocurrency pump-and-dumps,”
- [2] T. Li, D. Shin e B. Wang, “Cryptocurrency pump-and-dump schemes,”
- [3] F. Allen e D. Gale, “Stock price manipulation”
- [4] Mei, G. Wu, e C. Zhou, “Behavior based manipulation: theory and prosecution evidence,”
- [5] A. I. Khwaja e A. Mian, “Unchecked intermediaries: Price manipulation in an emerging stock market,”
- [6] L. Frieder e J. Zittrain, “Spam works: Evidence from stock touts and corresponding market activity,”
- [7] M. Siering, “All pump, no dump? the impact of internet deception on stock markets.”
- [8] N. Gandal, J. Hamrick, T. Moore e T. Oberman, “Price manipulation in the bitcoin ecosystem,”
- [9] P. M. Krafft, N. Della Penna e A. S. Pentland, “An experimental study of cryptocurrency market dynamics,”
- [10] M. La Morgia, A. Mei, F. Sassi e J. Stefa, “Pump and Dumps in the Bitcoin Era: Real Time Detection of Cryptocurrency Market Manipulations”
- [11] S. Asamadu Sarkodie, A. Maruf Yakubu, P. Asantewaa Owusu, “COVID-19 pandemic improves market signals of cryptocurrencies—evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin” (2022)
- [12] Shima Imani, Alireza Abdoli, ali Beyram, Azam Imani, Eamonn Keogh, “Multi-Window-Finder: Domain Agnostic Window Size for Time Series Data” (2021)

Sitografia

[1] Come si manipolano le criptovalute

<https://www.milanofinanza.it/news/pump-and-dump-cosi-si-manipolano-le-criptovalute-201808061936404179>

[2] Truffe effettuate sulle criptovalute

<https://www.cybersecurity360.it/nuove-minacce/le-frodi-tramite-criptovalute-come-riconoscerle-i-metodi-piu-comuni-e-come-difendersi/>

[3] Dati sulle truffe ai danni di proprietari di NFT

<https://collezionedatiffany.com/rubati-piu-di-100-milioni-di-in-nft-in-un-solo-anno>

[4] Il funzionamento degli schemi Pump and Dump.

<https://coinrivet.com/it/guides/security/what-is-a-pump-and-dump-crypto-scheme/>

[5] Definizione di blockchain.

<https://it.wikipedia.org/wiki/Blockchain>

[6] Le principali applicazioni della tecnologia blockchain.

<https://aws.amazon.com/it/what-is/blockchain/>

[7] Il funzionamento della Blockchain.

<https://www.borsaitaliana.it>

[8] Com'è nato il Bitcoin.

<https://www.money.it/Chi-e-Satoshi-Nakamoto-creatore-Bitcoin>

[9] Bitcoin e le altre criptovalute.

<https://www.geopop.it/bitcoin-e-le-altre-criptovalute-cosa-sono-e-come-facciamo-ad-attribuirgli-un-valore/>

[10] Il caso Gamestop

https://it.wikipedia.org/wiki/GameStop#Caso_boom_azioni_GameStop

[11] La vicenda della Stratton Oakmont

https://it.wikipedia.org/wiki/Stratton_Oakmont

[12] Il fallimento di Enron

<https://it.wikipedia.org/wiki/Enron>

[13] Telegram

<https://it.wikipedia.org/wiki/Telegram>

[14] Discord

<https://it.wikipedia.org/wiki/Discord>

[15] Reddit

<https://it.wikipedia.org/wiki/Reddit>

[16] BitcoinTalk

<https://bitcointalk.org/>

[17] PADL

<https://padl.mine.nu.siteindices.com/>

[18] 2018 Cryptocurrency bubble

https://en.wikipedia.org/wiki/Cryptocurrency_bubble#:~:text=By%20September%202018%2C%20cryptocurrencies%20collapsed,value%20in%20the%20previous%20week.

[19] Andamento del mercato delle criptovalute nel 2019

<https://it.cointelegraph.com/news/someone-say-bear-market-bitcoin-sets-3-huge-new-records-in-2019>

[20] Moving Standard Deviation formula

https://www2.microstrategy.com/producthelp/current/FunctionsRef/Content/FuncRef/MovingStDev__moving_standard_deviation_.htm

[21] Simple Moving Average formula

https://en.wikipedia.org/wiki/Moving_average

[22] Utilizzo degli indici a media mobile nelle analisi di mercato

<https://www.investopedia.com/articles/active-trading/052014/how-use-moving-average-buy-stocks.asp>

[23] Introduzione al Matrix Profile

<https://towardsdatascience.com/introduction-to-matrix-profiles-5568f3375d90>

[24] Definizione di Algoritmo Anytime

https://it.wikipedia.org/wiki/Algoritmo_anytime

[25] Algoritmo STOMP

https://matrixprofile.org/matrixprofiler/reference/mp_algos.html

Ringraziamenti

Ora che la parte tecnica è conclusa vorrei ringraziare alcune persone che hanno contribuito all'epilogo di questo percorso, spero di non dilungarmi troppo ma poiché non ho avuto la stessa possibilità dopo la Triennale non vorrei “bruciare” l'occasione.

Vorrei ringraziare chi non è stato citato nelle pagine precedenti ma ha svolto un lavoro fondamentale di coordinamento e mi ha saputo guidare lungo il sentiero (spesso molto tortuoso ed interrotto per colpa mia) di analisi e scrittura di questa tesi: il professor Luca Virgili.

In seguito, non per importanza ovviamente, vorrei ringraziare i miei genitori che mi hanno sempre sostenuto con tenacia e mi hanno sempre spronato a dare il meglio di me e a non mollare mai anche quando sembrava impossibile farcela. Senza i quali oggi non sarei qui e non probabilmente non sarei ciò che sono.

Vorrei ringraziare mia sorella che con un'innocente domanda: “Come va la tesi?” mi ricordava che il lavoro non era compiuto e andava portato al termine anche nei momenti in cui sembrava impossibile.

Vorrei ringraziare i miei zii che mi hanno sempre sostenuto e le quali parole a volte mi hanno dato la tenacia per continuare a fare ciò che stavo facendo.

Vorrei ringraziare anche coloro che non possono vedermi oggi ma che avrei tanto voluto avere con me, i miei nonni.

Vorrei ringraziare tutti miei amici, in particolare i “Daddy’s Children”, il mio amico di una vita Luca e gli “Snakes” (che da tanto aspettano il tanto atteso invito alla mia Laurea) e mi hanno incoraggiato ad andare avanti per concludere il mio percorso.

Infine, vorrei ringraziare i compagni di corso, con i quali si è creato un legame molto importante, sui quali ho potuto sempre contare e con i quali ho vissuto dei momenti che porterò con me per tutta la vita.