NATIONAL RESEARCH UNIVERSITY

HIGHER SCHOOL OF ECONOMICS

International College of Economics and Finance

Timur Pugoev

Cryptocurrency Pump-and-Dump Schemes, Detecting, Modelling and Forecasting

Master’s Thesis
38.04.01 ECONOMICS
Master’s Programme “**Financial Economics**”

Referee	Scientific advisor
Professor at ICEF,	Professor at Faculty of Economic Science,
Higher School of Economics	Higher School of Economics
Boulatov Alexei	Dean Fantazzini

Moscow 2023

Chapters

1. Introduction

Pump-and-dump schemes have long been a scourge in the world of finance, but with the rise of cryptocurrencies, these fraudulent activities have taken on new forms and become more prevalent than ever. Cryptocurrency pump-and-dump schemes represent a growing threat to investors in the emerging world of digital currencies. This phenomenon is not new to the financial world, but it has taken on new dimensions in the realm of digital currency. As more people become interested in investing in cryptocurrencies, the potential rewards for those who can manipulate the market become greater. As a result, detecting, modelling, and forecasting cryptocurrency pump-and-dump schemes has become a significant challenge for regulatory authorities and investors alike. To overcome this challenge and create appropriate methodology we have to understand that there are the huge differences between pump-and-dump schemes in the traditional economic context and the cryptocurrency context.

In the traditional economic context, pump-and-dump schemes often involve low-value securities, such as penny stocks. These securities are often associated with small companies that have little or no track record, making them vulnerable to manipulation by unscrupulous investors. The scheme typically involves a small group of people who coordinate their efforts to drive up the price of the stock by spreading misleading information, such as false news or rumors. Once the price has reached a certain level, the perpetrators sell their shares, causing the price to fall, and leaving other investors with significant losses.

In contrast, the cryptocurrency context is characterized by decentralization, anonymity, and a lack of regulatory oversight. Cryptocurrencies are traded on decentralized platforms that allow for fast and anonymous transactions, making it easier for perpetrators to manipulate the market. In addition, the widespread use of social media and online platforms has made it easier for perpetrators to spread misinformation and hype around a particular cryptocurrency. The schemes often involve coordinated efforts by groups of investors to buy up a particular cryptocurrency and drive up its price through online hype and social media promotion. Once the price has reached a certain level, the perpetrators will sell off their holdings, causing the price to crash and leaving other investors with significant losses.

Well, pump-and-dump schemes are fraudulent investment strategies that involve artificially inflating the price of an asset, followed by the perpetrators selling their shares at a profit before the price falls. P&Ds are not new, but they have evolved in the cryptocurrency context, where they are more prevalent and difficult to track due to the decentralized nature of cryptocurrencies. As a result, investors need to be cautious when investing in cryptocurrencies and avoid investment opportunities that promise quick, large returns with little or no risk.

Investors must conduct thorough research and seek advice from trusted sources before making any investment decisions. However, currently there is a few articles on this topic, that's why we must conduct a comprehensive study. So, the object of this research – P&D schemes, the subject – detection possibility of P&D schemes on cryptocurrency exchanges.

This research will also explore some of the key issues surrounding cryptocurrency pump-and-dump schemes, including the mechanics of how these schemes work, the methods that are commonly used to detect them, and the challenges that remain in predicting their occurrence. By understanding more about how these schemes work, and by developing effective models and strategies for detecting and mitigating their impact, we can help to protect investors and build a more stable and resilient financial system for the future.

To do so we must set the following task:

1. Identify and discover existing methods for detecting and forecasting the pump-and-dump schemes,
2. Analyze most resent methods to find anomalies in data and apply them to detect P&Ds,
3. Compare and interpret results of valuation using new and previous state of the art approaches.

In this work we created the unique dataset with 76 pumps that were collected for the last 2 years [2020-2022] from different Telegram channels and present relevant analysis to such data. We used different models to detect P&D schemes: Random Forest, SVM from the classical state-of-the-art Machine Learning models, and C-LSTM from the Deep Learning models. We found that, firstly, even if it becomes harder to detect pump-and-dump scheme, it is for sure possible. The Random Forest and SVM got the following results according to F1-score: 79.9%, 50% respectively. Secondly, we also show that Deep Learning models outperform classical Machine Learning models in a significant way of more than 5% in F1 metric.

The paper structure is the following: in Chapter 2 we discussed the history and background of pump-and-dump schemes, the start in the traditional economy, then in Chapter 3 we delve into crypto context and research the most cited and interesting papers, in Chapter 4 we analyzed different methods and models that can be used to detect P&Ds, in Chapter 5 we create the dataset and build announced models, and, finally, in Chapter 6 we outline the conclusion. All data and code can be found in the GitHub[1].

[1] https://github.com/TimurPugoev/pump-and-dump-detection-marterthesis

2. Background of P&D schemes, Traditional Economic Context

The stock market can be a complex and confusing world, and unfortunately, it is also a playground for fraudulent activities such as pump and dump schemes. These types of investment scams have been around for decades, but with the advancement of technology and the internet, they have become even more prevalent and sophisticated. To fully understand how pump and dump schemes work and what measures can be taken to prevent them, it is important to delve into the background and history of these fraudulent practices.

The origins of pump and dump schemes can be traced back to the early 20th century, when they were known as "bucket shops". These shops were essentially illegal gambling establishments that allowed customers to place bets on the direction of stock prices. The bucket shops would take the opposite side of these bets, and when the customer lost, they would keep the money. As the stock market became more regulated in the mid-20th century, bucket shops began to decline in popularity. However, the basic concept of manipulating stock prices for personal gain remained, and pump and dump schemes began to emerge in the 1980s. That’s all about traditional economic context.

As it has already mentioned, pump-and-dump schemes in traditional finance are a form of securities fraud that involves artificially inflating the price of a stock or other security through false or misleading information, then selling the security at the inflated price. This type of fraud is illegal and can have serious consequences for both the perpetrators and the victims. Check Figure 1 as an example of such scheme.

But how such misleading information spread now and a long time ago? What types of stock are usually used in P&D schemes in traditional markets? Let’s try to answer to these and other possible question related to traditional finance.

2.1 History

Pump-and-dump schemes have been around for centuries, but they have become more prevalent and sophisticated with the rise of modern financial markets. These schemes are often carried out by individuals or groups who are looking to make a quick profit by manipulating the market. This scheme typically involves the promotion of a low-priced stock through various channels, such as social media, email newsletters, or online forums. The promoters then create a buzz around the stock and drive-up demand, causing the price to rise rapidly.

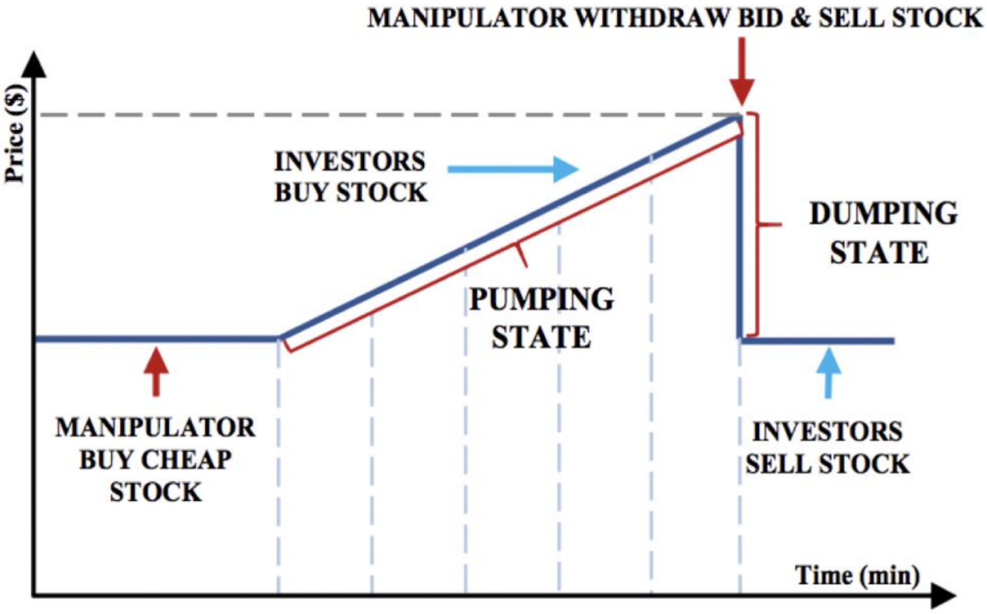


Figure 1. Pump and Dump scheme
Source: Ramos, 2017[2]

Once the stock has reached a certain price point, the promoters sell their holdings, causing the price to crash and leaving unsuspecting investors with significant losses. So, it’s important to be aware of the risks and to do your own research before making any investment decisions.

The pump and dump scheme is the basis for the plots of two popular movies: "Boiler Room" and "The Wolf of Wall Street". Both films depict an office full of stockbrokers who sell shares of dubious companies over the phone. The company itself, which also owns a large number of shares, encourages brokers with good bonuses for selling securities to as many customers as possible. The high demand for stocks drives up their price. Once the buyers dry up, the company dumps its shares at the peak price and makes a huge profit. This causes the price to collapse, resulting in losses for the clients. Originally, the scheme worked through phone calls, but now it has moved online: scheme creators send emails and publish fake news. Usually, this is presented in the form of insider information, such as an upcoming merger, a large order, or the approval of a company's product by a regulatory agency.

In recent years, the scheme has also worked through messaging apps and social media. As a result, the investor community can push the price in the desired direction. This is how meme stocks like Gamestop[3] have emerged. In January 2021, the company's shares prices soaring up to 800% in just one trading session without any significant reason such as M&A announce or financial statement publication. One of the most famous examples of a pump-and-

[2] «Cryptocurrency Pumping Predictions: A Novel Approach to Identifying Pump And Dump Schemes», Cameron Ramos, Noah Golub, 2017, Stanford students, Final project
[3] “GameStop Stock Skyrockets Again After NFT And Crypto Market Plans Emerge”, Forbes, Jan 6, 2022

dump scheme in the traditional economy is the South Sea Bubble of the early 18th century. The South Sea Company was created to trade with Spanish America, but its shares quickly became a speculative bubble, with investors bidding up the price of the stock in the hope of making a quick profit. When the bubble burst, many investors lost their life savings. There is one more notable example is the case of Enron, which used accounting fraud to inflate its stock price before ultimately collapsing in a massive scandal, we will cover all these cases later. All in all, there have been many examples of pump-and-dump schemes in the stock market.

In messaging apps and social networks, it is possible to organize a price increase for a particular asset. For example, in Telegram it's possible to spread free signals and make thus money on schemes, usually you have to pay to get into such groups. Members of paid groups enter the game during the accumulation phase when the price has not yet risen. Free groups notify about the pump at a later stage when the price has already increased. Also, there are several trial runs before the main impulse to test the ground: where resistance is possible, what is the overall mood of other holders, and so on.

2.2 Participants and Assets

The participants in pump-and-dump schemes can include individuals, groups, or even entire companies. The perpetrators typically have some level of knowledge or influence in the market, and they use this to their advantage to manipulate the price of a particular security.

Pump and dump schemes are usually carried out with so-called penny stocks - securities of small-cap companies often listed on over-the-counter markets. These markets are poorly regulated, have low liquidity, no market makers, and prices are easier to manipulate. Therefore, it does not require much money and buyers to push the quotes. In addition, such companies are less transparent and do not provide enough information about their activities. Thus, any news or rumor can be easily used by scammers, and potential investors do not have the opportunity to verify everything. In recent years, the pump and dump scheme has migrated virtually unchanged to the cryptocurrency market. This was facilitated by the lack of regulation. Currently, there are thousands of cryptocurrencies, many of which are little known and have low market capitalization. Their price can be easily pumped. Unlike the crypto market, there is a significant regulation in traditional economy. So, then we have to answer the question - how legal is the scheme?

2.3 Regulation

In the United States, until the 20th century, most market manipulations affected only a small circle of fairly affluent investors, and the government did not particularly interfere in these processes. The situation changed after World War I, when many Americans - middle-class representatives discovered the stock market. Manipulations became mass-scale, and in 1934, the Securities and Exchange Commission (SEC) was established to regulate the market.

This practice was eventually recognized as illegal. It is considered securities fraud, which can result in large fines and even imprisonment, depending on the scale of the scheme.

For example, in 1992, Meyer Blinder was sentenced to 46 months in prison for securities fraud. His company dealt with penny stocks and was known for the "three phone calls" method, where brokers would call clients multiple times before recommending a purchase to gain their trust. This tactic was used to sell stocks of shell companies, which allowed Blinder to earn $0.1 billion.

Another example is from September 2000 when the SEC investigated high school student Jonathan Lebed. He was the first minor to be held accountable for stock market fraud. For six months, he pumped stocks through the internet, posting hundreds of messages under fake names on the Yahoo-Finance message boards. He made anywhere from $12 thousands to $74 thousands per day. In the end, he had to return his illegally gained profits plus interest, totaling $285 thousands.

In 2021, John McAfee, the founder of a well-known antivirus software company, was accused of making millions by allegedly manipulating the market price of some cryptocurrencies.

SEC regulatory measures and FINRA - the Financial Industry Regulatory Authority - allow to some extent to curb the shady schemes of brokers like the Blinder case. However, this does not protect against pump and dump from individual citizens and unregistered groups like the community of users online. For example, in the case of the GameStop stock, the SEC conducted an investigation and found no evidence of fraudulent activity. In Russia, the regulator is the Central Bank, which from time-to-time records price manipulation in relation to third-tier stocks. But usually, it is limited to warnings that it cannot be done and a small fine. According to the Code of Administrative Offenses, the fine for manipulation is only 3 to 5 thousand rubles. But if there is criminal liability, for example, in the case of an organized group or particularly large income, then it is much more serious: organizers face up to 7 years in prison.

2.4 A Few Examples

Let’s start with a pump and dump scheme with stocks is the case of Enron, a large energy company that filed for bankruptcy in 2001 after it was revealed that the company had engaged in widespread accounting fraud. Enron's executives had artificially inflated the company's stock price through a variety of fraudulent accounting practices (they used lots of SPE to manipulate off-balance liabilities and assets), and many investors who had purchased Enron stock at its peak lost a significant amount of money when the company's stock price collapsed.

Another example is the case of Stratton Oakmont, a brokerage firm that engaged in a large-scale pump and dump scheme in the 1990s. The firm would purchase large blocks of shares in small, relatively unknown companies and then use high-pressure sales tactics to convince clients to purchase those shares. As the price of the shares increased due to the increased demand, Stratton Oakmont would sell its own shares at a profit, leaving many of its clients with worthless stocks.

More recently, there have been instances of pump and dump schemes involving so-called "meme stocks" such as GameStop and AMC Entertainment. In early 2021, a group of amateur investors on the online forum Reddit banded together to drive up the price of GameStop stock in what was essentially a coordinated pump and dump scheme. While some investors made significant profits, many others lost money when the stock price eventually crashed.

All examples can be found in Wiki[4].

2.5 Literature Review

There are not a lot articles you can find about P&D schemes in traditional economic context with stocks and other financial instruments. There are only few of them that cited more than 50 times according to Google Scholar.

Let’s start with «Stock manipulation and its effects: pump and dump versus stabilization»[5]. This study delves into the intricate world of stock price manipulation in the Taiwan stock markets, shedding light on the factors that contribute to such practices and their impact on market efficiency. The researchers have collected a unique dataset that enables them

[4] https://en.wikipedia.org/wiki/Pump_and_dump

[5] «Stock manipulation and its effects: pump and dump versus stabilization», Yu Chuan Huang, Yao Jen Cheng, 2015, Review of Quantitative Finance and Accounting, volume 44, pages 791–815

to analyze the patterns and characteristics of manipulated stocks, providing valuable insights. The findings of this study reveal that the firms most vulnerable to stock price manipulation tend to be small and suffer from weak corporate governance. The manipulations themselves often follow the notorious "pump-and-dump" strategy, causing temporary price impacts, increased volatility, and large trading volumes. While such manipulations can lead to short-term price continuation, they also result in long-term price reversals, thereby disrupting market efficiency. Interestingly, the study also found that the impact of stock manipulation depends on the fundamentals of the firm in question. Manipulated firms with positive fundamentals tend to weather the storm better than those with negative fundamentals, which experience a more severe impact on market efficiency. All in all, this study provides a nuanced understanding of stock price manipulation in the Taiwan stock markets, highlighting its complex nature and its potential repercussions for market participants.

Another interesting article is «Unchecked intermediaries: Price manipulation in an emerging stock market»[6]. This study provides a unique and comprehensive analysis of the costs associated with poor governance of market intermediaries, using trade level data from the Pakistan stock market. The findings reveal a significant disparity in the rates of return earned by brokers trading on their own behalf, compared to those earned by outside investors. Surprisingly, neither market timing nor liquidity provision could explain this profitability differential. Instead, the study uncovers strong evidence of a P&D scheme. The colluding brokers artificially raise prices and attract positive-feedback traders, and then exit the market once prices have risen, leaving other investors to suffer the ensuing price fall. These manipulation rents account for almost half of total broker earnings, providing a plausible explanation for the difficulties in implementing market reforms and the marginalization of emerging equity markets. The study sheds light on the hidden costs of poor governance, emphasizing the need for greater transparency and accountability in market intermediaries. The findings also suggest that policymakers and regulators must prioritize measures to combat manipulation practices to promote investor confidence and the growth of emerging equity markets.

[6] «Unchecked intermediaries: Price manipulation in an emerging stock market», Asim Ijaz Khwaja, Atif Mian, 2005, Journal of Financial Economics, volume 78, pages 203-241

3. P&D schemes in the Cryptocurrency Context

Pump-and-dump schemes in recent years have become prevalent in the world of cryptocurrency. As we already discussed, the concept of pump-and-dump schemes has been around for a long time, but it has become much easier to execute with the rise of social media and the Internet, especially Internet and its instant messaging (IM) services such as Telegram. The crypto market is particularly susceptible to these schemes due to its lack of regulation and high volatility.

3.1 History

In the world of cryptocurrencies, the explosive growth and lack of regulation have made it a prime target for fraudulent activities. Pump-and-dump schemes have become a common occurrence, often carried out by a single online group with a concerted effort to manipulate the market. Despite the growing concern surrounding regulatory hurdles, the crypto market remains largely unregulated, allowing fraudulent activities to flourish at an unprecedented level compared to the more heavily regulated stock market. As the crypto space continues to evolve, detecting and preventing fraudulent activities at scale remains a critical challenge.

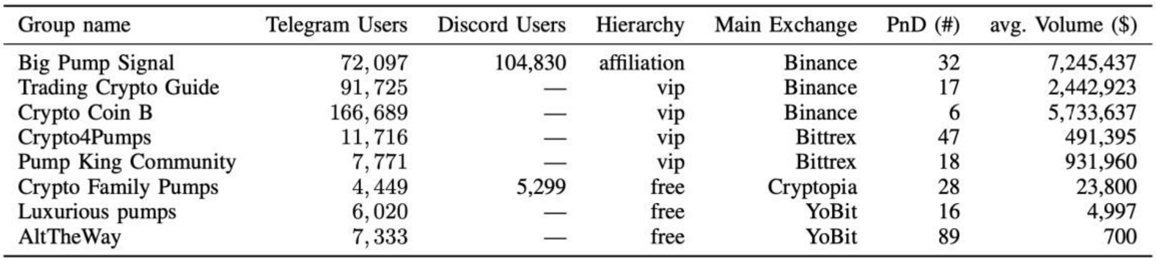
A pump in cryptocurrency market is a coordinated and intentional increase in the demand for a particular cryptocurrency/coin, resulting in a short-term price hike. The rise of encrypted chat applications such as Telegram and Discord has provided a breeding ground for various forms of misconduct in the cryptocurrency trading world.

3.2 Participants and Assets

There are 3 actors in such schemes: pump organizers, pump participants and pump target exchange. Individuals or groups who wish to execute pump-and-dump schemes utilize online to coordinate their activities, providing them with an unfair advantage through access to insider information. These individuals are referred to as pump organizers. In contrast, pump participants refer to traders who receive instructions the pump organizers and collectively purchase a specific coin, leading to an artificial price increase, or a "pump." However, many of these participants end up making no profit or even suffer losses due to the inflated prices. This highlights the inherent risks associated with pump-and-dump schemes, which can result in significant financial losses for unsuspecting investors. In addition to the pump organizers and participants, the pump target exchange also plays a crucial role in the scheme. Some exchanges are directly associated with pump-and-dump activities, with Yobit being a prime example of this. By organizing pumps on their platform, exchanges can profit by dumping coins they acquired before a pump at a higher price. In addition, they generate significant transaction fees as a result of the heightened trading volume that accompanies a P&Ds. Furthermore, these exchanges may take advantage of their early access to users' order data in order to engage in front-running during the frenzied buying and selling activity of a pump-and-dump. However, such activities are unethical and illegal, and regulators must take strict measures to protect innocent investors from these fraudulent schemes.

3.3 A Typical Pump-and-Dump Process

The pump-and-dump game is strong in the crypto world, and it all starts with a group chat or channel. Currently there are lots of such Telegram groups with significant number of followers. For example, you can check Table 1, it’s some metrics about channels, such data were collected by Massimo La Morgia et. al. for 2 years starting from 2017, we can see that there are some channels that have more than 100k subscribers. Besides, there are some channels that have 400.000 plus subscribers such as Binance Crypto Pumps[7].



Group name	Telegram Users	Discord Users	Hierarchy	Main Exchange	PnD (#)	avg. Volume ($)
Big Pump Signal	72,097	104,830	affiliation	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

Table 1. Metrics of Pump and Dump groups
Source: Massimo La Morgia, 2020[8]

So, there are several steps to create such groups. Firstly, pump organizers recruit as many subscribers as possible using different web-sites like Reddit, Bitcointalk, Steemit, usually they post invitation links. Once they hit over approximately one thousand members, the admins announce the next pump a few days ahead of time. They tell everyone the exact

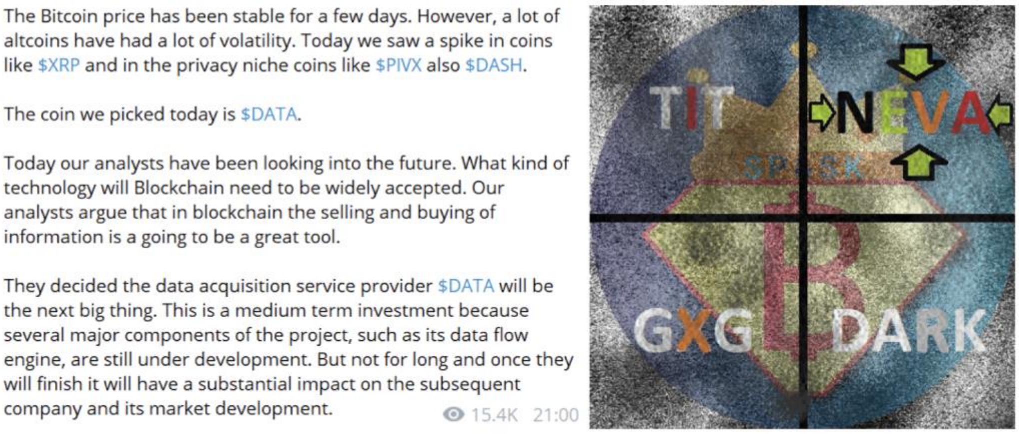
[7] https://t.me/Big_Pumps_Signals

[8] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

time and date of the announcement, where the pump will happen, and what coin to transfer funds into beforehand.

As the pump time approaches, the admins start a countdown and remind everyone of the "rules." It's all about buying fast, "shilling" the pumped coin to attract outsiders, and holding onto it for a bit to let others get in on the action. They even give pep talks and share stories of past successful pumps to get everyone hyped up.

Finally, the pump begins, and the admins announce the coin in a sneaky OCR-proof image to throw off the machines (that’s not always the case, sometimes it’s just a message with the name of the coin, check DATA coin on Figure 2). It's becoming increasingly common for cyber criminals to use obfuscation techniques to spread their malicious intent.



The Bitcoin price has been stable for a few days. However, a lot of altcoins have had a lot of volatility. Today we saw a spike in coins like $XRP and in the privacy niche coins like $PIVX also $DASH.

The coin we picked today is $DATA.

Today our analysts have been looking into the future. What kind of technology will Blockchain need to be widely accepted. Our analysts argue that in blockchain the selling and buying of information is a going to be a great tool.

They decided the data acquisition service provider $DATA will be the next big thing. This is a medium term investment because several major components of the project, such as its data flow engine, are still under development. But not for long and once they will finish it will have a substantial impact on the subsequent company and its market development.

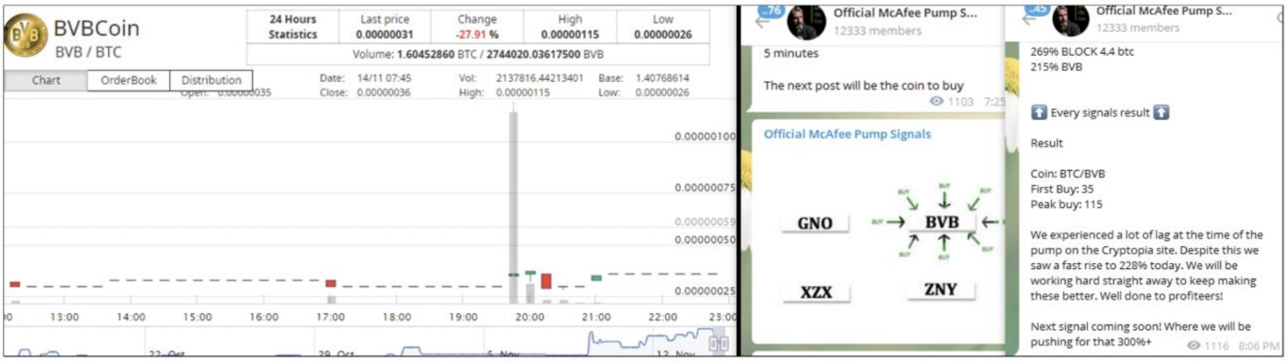
Figure 2. Two different types of messages to start of a pump and dump scheme on the DATA coin (left), the NevaCoin (right).
Source: Massimo La Morgia, 2020

For instance, a message that instructs the start of a pump and dump operation on a cryptocurrency like NevaCoin (Figure 2.) may be disguised to impede bots from parsing it with OCR methods and initiating market operations quicker than humans. While such actions could cause significant harm, it's critical to avoid engaging with malicious content and to stay vigilant against potential cyber threats.

After the announcement the coin price surges like crazy for the first minute, but then, just as quickly, it starts to fall. The admins shout to "BUY BUY" and "HOLD HOLD," but the damage is already done. Panic-selling begins, and the price drops back down, sometimes even lower than before.

Within half an hour, the admins post a review of the pump, usually only including the start and peak prices to make it look like the profits were huge. They conveniently leave out any information about low trading volume or short time frames. So, while it might seem like a fun game to join in on, remember that you're playing with fire and might end up getting burned.

In Figure 3 you can check the real successful P&D event. On the right-hand side of the screenshot is the message history of a Telegram channel. The message history of a Telegram channel is displayed on the right-hand side of the screenshot, detailing the final countdown, coin announcement, and pump result. Meanwhile, the left-hand side of the screenshot depicts the market movement of the relevant cryptocurrency during the corresponding pump period.



BVBCoin
BVB / BTC

24 Hours Statistics	Last price	Change	High	Low
	0.00000031	-27.91 %	0.00000115	0.00000026

Volume: 1.60452860 BTC / 2744020.03617500 BVB

Chart | OrderBook | Distribution

Open: 0.00000035 Date: 14/11 07:45 Vol: 2137816.44213401 Base: 1.40768614
Close: 0.00000036 High: 0.00000115 Low: 0.00000026

Official McAfee Pump S...
12333 members

5 minutes

The next post will be the coin to buy
1103 7:25

Official McAfee Pump Signals

Official McAfee Pump S...
12333 members

269% BLOCK 4.4 btc
215% BVB

⬆ Every signals result ⬆

Result

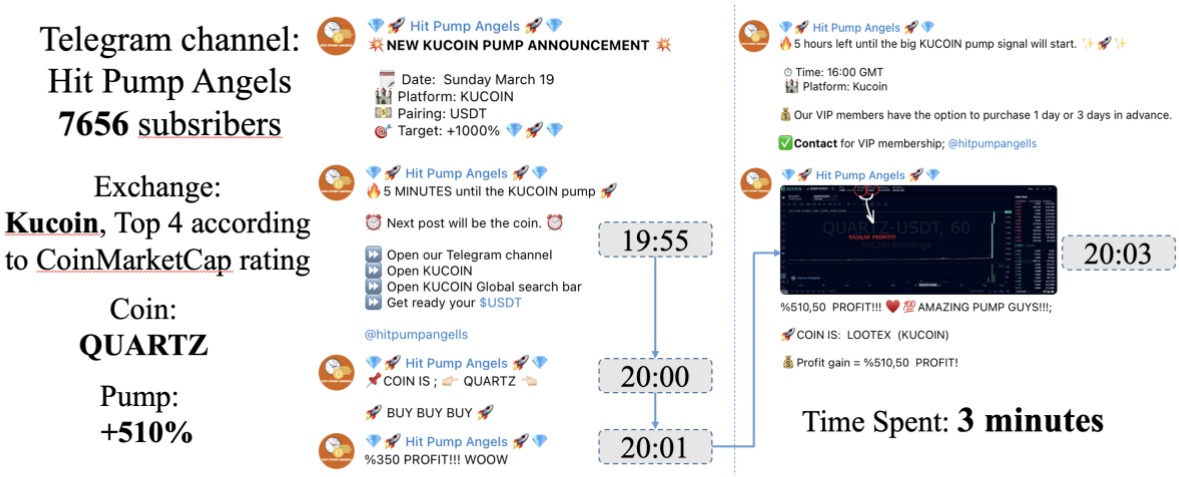
Coin: BTC/BVB
First Buy: 35
Peak buy: 115

We experienced a lot of lag at the time of the pump on the Cryptopia site. Despite this we saw a fast rise to 228% today. We will be working hard straight away to keep making these better. Well done to profiteers!

Next signal coming soon! Where we will be pushing for that 300%+
1116 8:06 PM

Figure 3. Successful Pump and Dump event
Source: J Xu, 2019[9]

It also important to say that there is a huge difference in time scale of pump and dump scheme comparing to traditional one. Usually, the pump in traditional exchange can last for a month or more, but that's not the case with P&D schemes in crypto market context. Here just an example of a time scale:



Telegram channel:
Hit Pump Angels
7656 subsribers

Exchange:
Kucoin, Top 4 according to CoinMarketCap rating

Coin:
QUARTZ

Pump:
+510%

💎🚀 Hit Pump Angels 🚀💎
💥NEW KUCOIN PUMP ANNOUNCEMENT 💥

🗓 Date: Sunday March 19
🏰 Platform: KUCOIN
💵 Pairing: USDT
🎯 Target: +1000% 💎🚀💎

💎🚀 Hit Pump Angels 🚀💎
🔥5 MINUTES until the KUCOIN pump 🚀

⏰ Next post will be the coin. ⏰

⏩ Open our Telegram channel
⏩ Open KUCOIN
⏩ Open KUCOIN Global search bar
⏩ Get ready your $USDT

@hitpumpangells

💎🚀 Hit Pump Angels 🚀💎
📌COIN IS ; 👉 QUARTZ 👈

🚀 BUY BUY BUY 🚀

💎🚀 Hit Pump Angels 🚀💎
%350 PROFIT!!! WOOW

💎🚀 Hit Pump Angels 🚀💎
🔥5 hours left until the big KUCOIN pump signal will start. ✨🚀✨

⏱ Time: 16:00 GMT
🏰 Platform: Kucoin

💰 Our VIP members have the option to purchase 1 day or 3 days in advance.

✅**Contact** for VIP membership; @hitpumpangells

💎🚀 Hit Pump Angels 🚀💎

%510,50 PROFIT!!! ❤️ 💯 AMAZING PUMP GUYS!!!;

🚀COIN IS: LOOTEX (KUCOIN)

💰 Profit gain = %510,50 PROFIT!

Time Spent: **3 minutes**

Figure 4. Time scale of P&D scheme on crypto market
Source: Own analysis

[9] "The Anatomy of a Cryptocurrency Pump-and-Dump Scheme", J Xu, B Livshits, 2019, USENIX Security Symposium

It lasts for a few minutes, that' actually the case for such schemes, some of them may last for a minute, some of them ten plus, but not more than a day or a week.

Besides, it should be noted that not every P&D attempt is successful. Usually that's because of the unanticipated price movement of the target coin, and it is unknown what cause such shifts. An alternate theory could be that someone had discerned the pattern for coin selection and procured a batch of coins with a greater potential for pump, including the specific coin in question, prior to the pump. This could clarify the anomalous movements observed by the admin.

3.4 A Few Examples

In recent years, there have been numerous examples of pump-and-dump schemes in crypto, including the "Wolf of Wall Street" pump group, which was shut down by the SEC in 2018 after defrauding investors out of millions of dollars, and the "Shitcoin Trading" group, which was shut down in 2019 after running a pump-and-dump scheme on various altcoins. Additionally, several individual cryptocurrencies have been subject to pump-and-dump schemes, such as Bitconnect, which collapsed in 2018 after being exposed as a Ponzi scheme.

One of the most infamous cases occurred in 2018, when the price of the little-known cryptocurrency, Viacoin, surged by over 200% in just a few hours. This sudden price hike was due to a coordinated pump and dump scheme orchestrated by a group of individuals on the Binance exchange.

Another example happened in 2021, when the cryptocurrency Safemoon experienced a dramatic increase in price, fueled by social media hype and promotion by influencers. However, the price soon plummeted, leaving many investors with significant losses. The Safemoon case highlights the danger of blindly following hype and promises of quick profits in the volatile world of cryptocurrency trading.

In another recent case, the US Securities and Exchange Commission (SEC) charged a group of individuals with operating a P&D scheme for two cryptocurrencies, Bitcoiin2Gen and Ethereum Meta. The group allegedly used false and misleading statements to inflate the price of these cryptocurrencies, before dumping their own holdings for a significant profit.

These examples demonstrate the pervasive nature of pump-and-dump schemes in the world of cryptocurrency, and the need for greater regulation to protect investors from fraudulent activities.

3.5 Regulation in in the Cryptocurrency Context

Crypto coins, a digital or virtual currency that uses cryptography for security, has taken the world by storm in recent years. Despite the many benefits that come with cryptocurrency, such as decentralization and anonymity, governments are still grappling with how to regulate this emerging technology. In this chapter, we will explore the various regulations that have been put in place to govern the cryptocurrency market.

One of the primary issues that regulators face is the lack of uniformity in the definition of cryptocurrency. Different countries define cryptocurrency in different ways. For example, the United States defines cryptocurrency as a digital asset that can be used as a medium of exchange, while in Japan, cryptocurrency is recognized as a legal payment method. This lack of uniformity has resulted in inconsistent rules and regulations governing cryptocurrency. To regulate the cryptocurrency market, governments have implemented several strategies, including licensing and registration, taxation, and surveillance.

Licensing and registration ensure that cryptocurrency exchanges and service providers adhere to a set of rules and regulations designed to protect consumers. The process involves obtaining a license from a regulatory agency, such as the Financial Conduct Authority (FCA) in the UK or the Securities and Exchange Commission (SEC) in the US.

Taxation is another strategy that governments use to regulate cryptocurrencies. Just like any other asset, cryptocurrency is subject to capital gains tax in most countries. For example, in the US, cryptocurrency is treated as property for tax purposes and is subject to the capital gains tax. In the UK, cryptocurrency is subject to income tax or capital gains tax, depending on how it is traded or used.

Finally, surveillance is essential to detect and prevent fraudulent activities such as money laundering, terrorism financing, and cybercrime. Governments have established agencies such as the Financial Action Task Force (FATF) to provide guidance for effective surveillance in the cryptocurrency market.

In conclusion, regulation in the cryptocurrency context remains a complex and evolving issue. Governments continue to explore different strategies to govern this emerging technology in a manner that protects the interests of both consumers and the wider economy. As the cryptocurrency market continues to grow and mature, it is important for regulators to keep pace and adapt their strategies accordingly. Ultimately, a well-regulated cryptocurrency market can encourage innovation and growth while minimizing risks for all stakeholders.

3.6 Literature Review, Empirical Analysis

So, now we must find out what models can help us to predict, detect and prevent pump-and-dump schemes, we mainly will focus on the most cited and interesting articles that used different state-of-the-art models such as Random Forest, LSTM, etc., however, we will also research literature about crypto market at all which can help us to delve into the theme. It must notice that the use of deep learning models in the detection of crypto fraud is a rapidly evolving area of research. By utilizing Anomaly Detection models via neural networks, newest articles demonstrates that deep learning methods can achieve state-of-the-art performance on the available data, outperforming classical machine learning and statistical models as RF (random forest), but we will talk about it later.

We will start exploring and researching well-known articles about cryptocurrencies and crypto markets manipulations. The article that is called «Price manipulation in the bitcoin ecosystem»[10] provide proof that the initial surge of Bitcoin's value to $1000 was a result of market manipulation. By analyzing the widely-used dataset of the Mt. Gox exchange, researchers uncovered fraudulent transactions conducted by two individuals known as the 'Willy bot' and 'Markus bot'. These bots were deliberately purchasing Bitcoin in order to artificially inflate the price and trading volume. There are some other results of this article:

- Around 80 percent of the days when suspicious trading activity took place saw an increase in prices. However, on approximately 55 percent of the days with no suspicious activity, prices rose as well.
- On days when the key actor was active (the bots), the USD/BTC exchange rate increased by an average of over $20 per day. In contrast, the exchange rate remained largely unchanged on the days when the actor was inactive.

Another article «An experimental study of cryptocurrency market dynamics»[11] conduct a unique analysis on the behavioral tendencies exhibited by users in the Cryptsy exchange market. Through their research, it is evident that the market can be influenced by even the smallest of buy transactions. To investigate this, they implement bots that carried out over 100,000 trades in 217 cryptocurrencies, each costing less than a penny, over the course of six month, these bots randomly purchase small amounts of various digital currencies, concluding

[10] «Price manipulation in the bitcoin ecosystem», N. Gandal, J. Hamrick, T. Moore, T. Oberman, 2018, Journal of Monetary Economics, vol. 95, pp. 86–96.

[11] «An experimental study of cryptocurrency market dynamics» P. M. Krafft, N. Della Penna, A. S. Pentland, in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 2018, p. 605.

that traders exhibit a preference for currencies with recent activity. Their results demonstrate that individual "buy" actions caused significant short-term increases in subsequent buy-side activity, exceeding the size of our interventions by hundreds of times. Authors also noted that the design choices made by the exchange they studied may have played a role in facilitating these and other peer influence effects, underscoring the potential social and economic significance of HCI in the design of digital institutions.

Authors of «A New Wolf in Town? Pump-and-Dump Manipulation in Cryptocurrency Markets»[12] showcasing how excessive self-assurance and a proclivity towards gambling can account for involvement in such schemes. By examining 355 instances over a six-month period, they have obtained compelling evidence to support both of these explanations. According to the article the distinctiveness of cryptocurrency pump-and-dumps poses two intriguing inquiries. Firstly, what motivates people to take part in these pumps? Secondly, how do manipulators gain a profit without deceiving participants or exploiting price asymmetry? To address these questions, they present a simple theoretical framework. Their research indicates that logical individuals, lacking an advantage in skill or speed, would not participate in pump-and-dump manipulations. This is because pumps yield negative expected returns for participants, other than manipulators, who can buy before the pump signal is issued. The reason for this is that pumps settle as a zero-sum game, redistributing wealth between players. Given that manipulators utilize their advantage to extract profits and trading costs, pumps evolve as a negative-sum game for non-manipulators. While faster or more skilled players can profit at the expense of slower or less skilled participants, collectively, non-manipulators lose money, presenting the puzzle of how these pumps can maintain participation. The pumps create significant market distortions, with an average of 65% deviation of price, and result in exceptionally high trading volumes in the millions of dollars (13.5 times the average volume), while also resulting in significant wealth transfers among participants.

We also must research «The Economics of Cryptocurrency Pump and Dump Schemes»[13]. It is one of the first articles that evaluates the extent of cryptocurrency pump and dump activities occurring on Discord and Telegram. In a six-month period during 2018, authors identified 3,767 different pump signals advertised on Telegram and another 1,051 on Discord. These schemes were promoting over 300 cryptocurrencies, revealing the widespread and profitable nature of this phenomenon. They further investigated factors that influence the

[12] «A New Wolf in Town? Pump-and-Dump Manipulation in Cryptocurrency Markets», Dhawan, Anirudh and Putnins, Talis J, 2021, Review of Finance, Forthcoming

[13] «The Economics of Cryptocurrency Pump and Dump Schemes», Feder, Amir & Gandal, Neil & Hamrick, JT & Moore, Tyler & Mukherjee, Arghya & Rouhi, Farhang & Vasek, Marie, 2018, CEPR Discussion Papers 13404, C.E.P.R. Discussion Papers

"success" of a pump, measured by the percentage increase in price near the pump signal. Authors discovered that the coin's rank, based on market capitalization and volume, is the most significant factor in determining the profitability of a pump. Pumping obscure coins, with lower volume, proves to be more profitable than the major coins in the cryptocurrency ecosystem.

Well, let's continue and deep into empirical literature review with articles that used classical machine learning approach to detect and forecast P&Ds. «The Anatomy of a Cryptocurrency Pump-and-Dump Scheme»[14], one of the most useful and interesting articles. The authors present a case study of a pump-and-dump schemes, investigate 412 P&D events that were successfully organized in Telegram groups from June, 2018 to February, 2019, and discover patterns in crypto-markets associated with such schemes.

They got very alluring results according to their market research. Firstly, they note that pump-and-dump activities can cause significant damage to the market. An analysis compared the three-hour trading volume of BTC for pumped coins before and during a pump-and-dump. The results show (check Figure 5.) that the artificially generated trading volume during the pump-and-dump period is astonishing: 8,793 BTC, primarily from Binance, is equivalent to approximately 50 million USD. This is nine times the pre-pump volume of 943 BTC in just eight months.

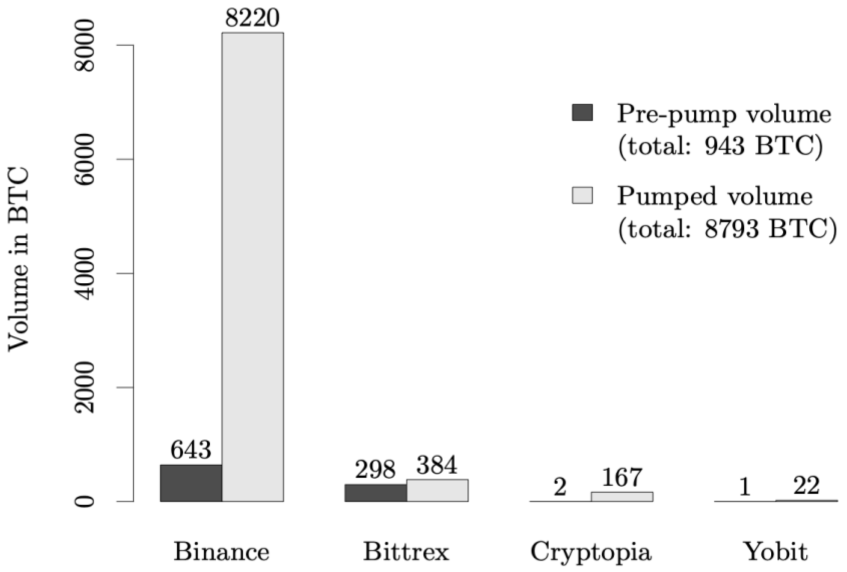
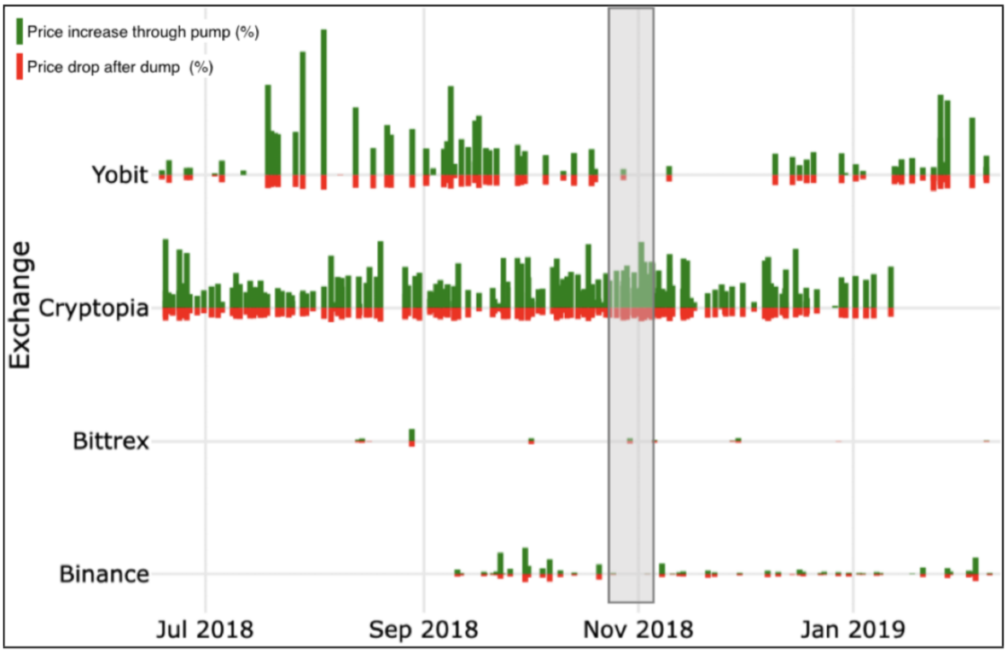


Figure 5. Aggregate Trading Volume of pumped crypto coins before and during a P&D scheme
Source: J Xu, 2019

[14] «The Anatomy of a Cryptocurrency Pump-and-Dump Scheme», J Xu, B Livshits, 2019, USENIX Security Symposium

Secondly, authors found the frequency and effectiveness of pump-and-dump activities conducted by individuals and displayed it in a graph (check Figure 6.). Bittrex is the least favored exchange in terms of frequency, while Binance is still gaining momentum since September 2018 but still encounters fewer instances of pump-and-dump compared to Yobit and Cryptopia. Yobit and Cryptopia, on the other hand, complement each other. Cryptopia experienced higher traffic when Yobit was inactive (especially from October 2018 to January 2019), and Yobit regained its popularity when Cryptopia went silent (since the hack in mid-January 2019). When examining the percentage increase in coin price during pump-and-dump schemes, it appears that Yobit and Cryptopia exchanges exhibit more significant increases compared to Bittrex and Binance. Additionally, the dump phase of these schemes results in a more severe drop in coin prices on Yobit and Cryptopia than on their counterparts.



(a) Pump and dump activities from June 2018 to February 2019

Figure 6. Pump and dump timeline[15]. All prices are denominated close price in BTC, and from a 3-hour window around pump activities.
Source: J Xu, 2019

Thirdly, activity distribution by exchange is the following:

- Among the 412 P&Ds (check Figure 7.), 21 (or 5%) took place in Bittrex exchange, 68 (or 17%) in Binance exchange, 112 (or 27%) in Yobit, and 211 (or 51%) in Cryptopia

[15] Visit http: //rpubs.com/xujiahuayz/pd for the full, interactive chart.

(only this exchange went into liquidation in May 2019[16]). To sum up, 146 out of 412 (or 35%) of the time, the selected coin had previously been pumped in the same exchange. Such result coincides with the result of the next article that we will research, we will talk about it later.

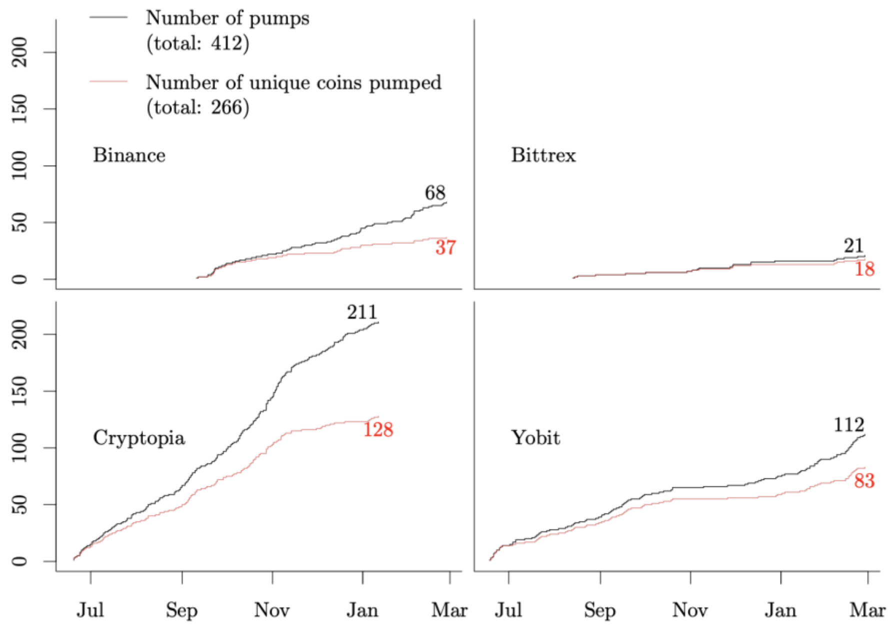


Figure 7. Cumulative number of P&Ds and number of unique pumped coins on four exchanges from June 2018 to February 2019
Source: J Xu, 2019

After conducting thorough market research, they have discovered that the movements in the market leading up to a pump-and-dump event can often provide clues as to which coin will be pumped. By utilizing LASSO regularized GML and balanced RF, several models have been created by authors to forecast the timing and location of a pump-and-dump announcement within a Telegram group. These models have demonstrated strong performance across all subsets (using the ROC-AUC metric), implying that it is feasible to anticipate which coins will be pumped using market data.

In addition, authors created highly effective trading strategy that can be used in conjunction with our prediction models. Their pre-purchase process involves buying a coin at the open price one hour prior to the coin announcement for each respective coin. To determine the amount of BTC to invest, they multiply the vote received from the random forest model by a constant, denoted as k. This ensures that the investment made on each coin is proportional to

[16] You can read more about it on this website: https://www.nasdaq.com/articles/cryptopia-exchange-currently-in-liquidation-gets-hacked-again%3A-report-2021-02-20

its vote, thereby increasing the investment on coins that have a higher likelihood of experiencing a pump. This approach is logical as a higher vote indicates a greater probability of a pump, thereby justifying a larger investment. Results from out-of-sample tests indicate that this strategy can yield returns of up to 60% over two and a half months.

In «Profitability of cryptocurrency Pump and Dump schemes» [17] author Tsuchiya, T conducted research to identify the features of P&Ds that were organized in Telegram and examine the market resilience to such activities. The study employed a Bayesian hierarchical framework to establish a regression model that illuminates the variables that impact the profitability (i.e., price change) of P&D attempts. Results unveiled that the impact of trading volume on profitability varies significantly across exchange markets. Specifically, Yobit and Cryptopia are more susceptible (easily tampered with) to trading volume increases than Binance and Bittrex, even when controlling for relevant factors such as timing of the pump (hourly, yearly), currency, and Telegram channel. Additionally, a machine learning model was developed to accurately identify successful schemes (price hikes) using pre-start information, achieving over 75% accuracy through tree-based ensemble models. This paper's contribution lies in its unique statistical approach, providing a comprehensive analysis of P&D schemes, with particular emphasis placed on the effect of each exchange. This sheds light on how social media groups manipulate the market, thus providing a better understanding of the market's workings.

We will continue our review with the article: « To the moon: defining and detecting cryptocurrency pump-and-dumps»[18], a first attempt to detect P&Ds using an adaptive threshold. This research paper analyses the literature on pump-and-dump schemes from traditional economic sources, connects it with cryptocurrencies, and proposes a set of standards to define a cryptocurrency pump-and-dump. These patterns of pump-and-dump display unusual behavior, therefore, techniques from anomaly detection research are utilized to identify instances of unusual trading activities to highlight potential pump-and-dump incidents. The results indicate that there are certain indications in the trading data that could aid in identifying pump-and-dump schemes. By examining a few real-world cases in their detection system, they demonstrate these signals.

The main focus of this paper is unsupervised anomaly detection as there was no labeled training data available at that time for cryptocurrency pump-and-dump schemes (in the current research we will use supervised learning approaches, because we will create such training

[17] «Profitability of cryptocurrency Pump and Dump schemes» Tsuchiya, T., 2021, Digit Finance 3, 149–167

[18] «To the moon: defining and detecting cryptocurrency pump-and-dumps», J. Kamps, B. Kleinberg, 2018, Crime Science

data). Conditional anomalies take into account the contextual information about the situation. Indicator variables and environmental variables are used to describe the contextual information. The values of the indicator variables may directly indicate an anomaly, while environmental variables may not. In the present context, the goal of authors was to identify the breakout indicators. The researchers have developed a real-time detector to identify pump and dump schemes using 1-hour candlesticks as input data. The detection process for pump-and-dump schemes can take up to an hour, though it is typically completed within 30 minutes. To detect these schemes, researchers utilize two anomaly thresholds based on transaction volume and coin price. These thresholds are determined by analyzing average windows of recent candlestick history. If both the price and volume surpass the computed thresholds, the event is classified as a pump-and-dump. Three different parameter configurations are provided by the researchers for threshold computation: Initial, Balanced, and Strict. The Basic configuration maximizes recall, the Strict configuration maximizes precision, while the Balanced configuration represents a trade-off between the two. However, the researchers were unable to provide accuracy scores in terms of precision and recall due to the lack of ground truth data in their dataset. According to the results they illustrate two cases where their system (with the Balanced parameter set) successfully detected a confirmed pump and dump scheme, and two cases where our system couldn't clearly identify the P&D. There are some articles that used such technique as a baseline in their own research and, without any doubt, outperform this threshold according to some metric.

Pump-and-dump schemes pose a significant challenge for crime science. Authors research has shown evidence of clustering in the data, with low market cap coins being targeted more frequently than higher market cap coins. This clustering can be utilized for preventative purposes by implementing strategies to mitigate potentially nefarious activity. A large majority of the cryptocurrency coins have a low market cap, with the top ten coins accounting for 85% of the overall market cap. Additionally, analysis of pump-and-dump schemes shows that around 30% of the symbols were responsible for approximately 80% of the pumps, suggesting that certain coins are targeted more frequently than others. This pattern reflects the concept of repeat victimization in environmental criminology literature. If a pump-and-dump chat group successfully targets a coin in the past, they may be more likely to pump that coin again (the same result according to the previous article). This can be used to prevent future nefarious activity by focusing efforts on these clusters, identifying why they are attractive targets, and implementing preventative strategies. An example of this clustering was observed by the authors in the case of Moonlight Signal (currently such Telegram group does not exist or it was renamed), who targeted the same coin (RDN[19]) twice within a two-week period.

Situational crime prevention ideas, such as increasing the effort required to commit a pump-and-dump, could prove useful in prevention efforts. If an exchange will require additional verification for investors to trade usually targeted symbol pairs which are determined to be vulnerable (shortly, such new approach would increase the effort required to trade), then it will save some inexperienced investors from such hazardous P&D schemes.

Assisting in preventing pump-and-dump schemes in the cryptocurrency world is really a challenging task. Cooperation is required between private entities such as cryptocurrency exchanges and government bodies to create effective prevention measures. Governments have allocated more resources to prevent such schemes, but exchanges may not be motivated to cooperate since they profit from trading activity. Increasing government regulation can also potentially undermine the decentralized nature of cryptocurrency trading. A solution could be interdisciplinary efforts from both practitioners and the research community to mitigate pump-and-dump schemes in the cryptocurrency world.

Let’s analyze «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations»[20]. In this study, authors discuss the phenomenon of "pump and dump" in the cryptocurrency ecosystem and present two relevant case studies. In the first study, they analyzed the gathered pump and dumps that were organized in Telegram channels on four different exchanges over time: Binance, Cryptopia, YoBit, and Bittrex. Throughout their longitudinal study spanning from July 2017 to January 2019, authors participated in over 100 groups and closely monitored their activities on a daily basis. Building membership and rapport with these groups granted them unique insights, including internal organizational hierarchies, pump and dump tactics, and strategies for recruiting external investors within the market. Table I outlines several metrics and distinguishing features of eight groups that they joined, which serve as representative examples. In the second study conducted by Massimo La Morgia et al., the focus was on Big Pump Signal, the largest group found during the research, which operates on Binance and has the capability of generating a transaction volume of 5176 BTC in a single operation. This is a higher volume than the combined total of 534 BTC generated by all the pump-and-dump schemes on Cryptopia, YoBit, and Bittrex. The authors of the study also introduced a novel real-time detection algorithm that is based on the anomalous growth of market buy orders, which are used when investors want to buy quickly

[19] https://coinmarketcap.com/currencies/raiden-network-token/

[20] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

at any price. According to their own tests, this algorithm outperforms the state-of-the-art methods for detecting pump-and-dump schemes according to their own times (it was unsupervised anomaly detection approach through indicator variables that was created and introduced by J. Kamps, B Kleinberg[21], we have already analyzed this article) in terms of both detection speed (reducing it from 30 minutes to 25 seconds) and F1-score (improving it from 60.5% to 92%).

In their research they focused only on Binance exchange, there are 2 main reasons: firstly, it's easier to collect data from Binance (not all exchanges give an opportunity to collect the historical data using ccxt[22] library that was used by the authors, only few hours before the current time are available), secondly, pump-and-dump schemes executed on other exchanges are orchestrated by groups with limited subscribers and overall views, which restricts their ability to target popular coins with high transaction volumes. As a result, they tend to focus on alt-coins with exceptionally low transaction volumes. Therefore, detecting pump-and-dump schemes on Binance is considered to be the most challenging and compelling task, as Binance is a popular exchange with a high trading volume and a vast array of coins. So, authors were able to collect 104 pump and dump events that were organized by 12 distinct groups. To gather data for these events, they collected trading data for a total of 14 days - 7 days prior to the event and 7 days after the event. In cases where multiple pump and dump events occurred within a short period of time for the same alt-coin, they just eliminated duplicate days from analysis. Ultimately, authors were able to gather trade records, including volume, price, operation type (buy or sell), and UNIX timestamps, for approximately 900 trading days, then they divide data on the following chunks: 5 seconds, 15 seconds, 25 seconds.

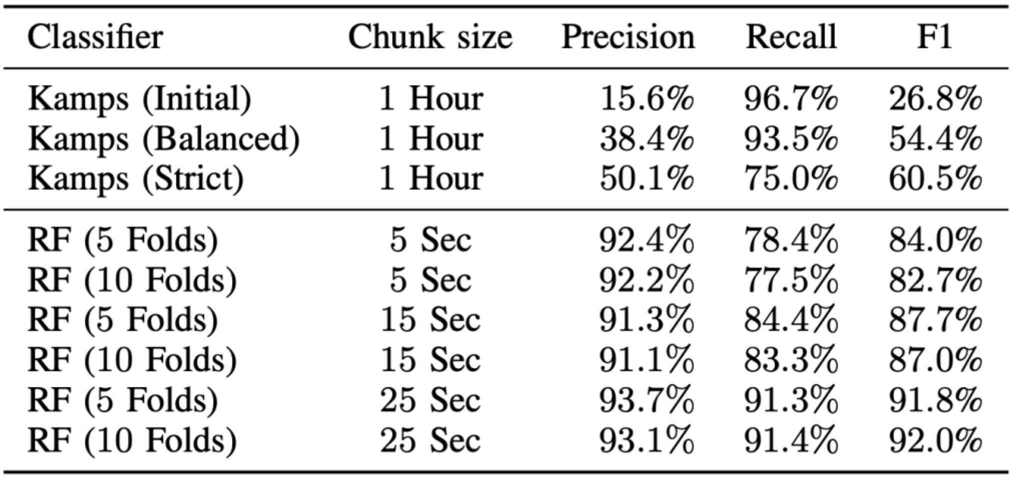
To analyze this data they used two different approaches: Random Forest and Logistic Regression. Authors find that Random Forest outperform Logistic Regression according to F1 score in each type of chunks. Finally, they report key results in Table (check Table 2.). As it was already mentioned, Random Forest outperform previous state-of-the-art model in terms of both detection speed and F1-score. Furthermore, the Random Forest classifier results highlight a correlation between chunk size and classifier performance. Despite precision remaining relatively stable across all time frames, recall improves with increasing chunk size dimensions.

Despite the fact that classical machine learning models can show good performance, currently the best models to predict anomaly such as P&D events are deep learning models according to F1 score or ROC-AUC or actually any other metric. So, «Crypto Pump and Dump

[21] «To the moon: defining and detecting cryptocurrency pump-and-dumps», J. Kamps, B. Kleinberg, 2018, Crime Science

[22] https://github.com/ccxt/ccxt

Detection via Deep Learning Techniques»[23] is appropriate article to prove the fact of outperforming of DL models. The authors propose 2 different models: C-LSTM, and the Anomaly Transformer. Their results show that LSTMs and Transformers have the ability to outperform easily both classical machine learning models and statistical models on this dataset.



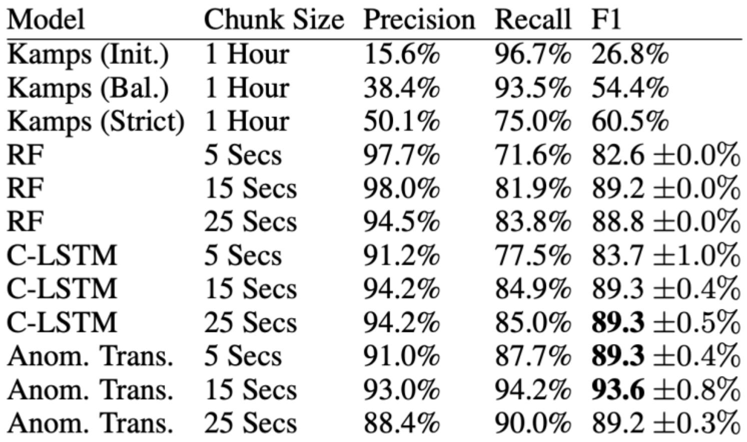
Classifier	Chunk size	Precision	Recall	F1
Kamps (Initial)	1 Hour	15.6%	96.7%	26.8%
Kamps (Balanced)	1 Hour	38.4%	93.5%	54.4%
Kamps (Strict)	1 Hour	50.1%	75.0%	60.5%
RF (5 Folds)	5 Sec	92.4%	78.4%	84.0%
RF (10 Folds)	5 Sec	92.2%	77.5%	82.7%
RF (5 Folds)	15 Sec	91.3%	84.4%	87.7%
RF (10 Folds)	15 Sec	91.1%	83.3%	87.0%
RF (5 Folds)	25 Sec	93.7%	91.3%	91.8%
RF (10 Folds)	25 Sec	93.1%	91.4%	92.0%

Table 2. Key model results: Kamps, Random Forest
Source: Massimo La Morgia, 2020[24]

The dataset utilized in this paper comprises of manually-labeled, unprocessed transaction data, which was originally curated by Massimo La Morgia et al., 2020, as previously mentioned. The deep learning models employed in this study demonstrated remarkable efficacy, surpassing all previous classical and statistical techniques, with the exception of the C-LSTM model on the 15-second chunked dataset, which yielded comparable outcomes. These results conclusively establish the potential of deep learning in an unexplored field. Notably, state-of-the-art results across all models are highlighted in Table 3.

[23] «Crypto Pump and Dump Detection via Deep Learning Techniques», Chadalapaka Viswanath, Chang Kyle, Mahajan Gireesh, Vasil Anuj, 2022, arXiv

[24] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)



Model	Chunk Size	Precision	Recall	F1
Kamps (Init.)	1 Hour	15.6%	96.7%	26.8%
Kamps (Bal.)	1 Hour	38.4%	93.5%	54.4%
Kamps (Strict)	1 Hour	50.1%	75.0%	60.5%
RF	5 Secs	97.7%	71.6%	82.6 ±0.0%
RF	15 Secs	98.0%	81.9%	89.2 ±0.0%
RF	25 Secs	94.5%	83.8%	88.8 ±0.0%
C-LSTM	5 Secs	91.2%	77.5%	83.7 ±1.0%
C-LSTM	15 Secs	94.2%	84.9%	89.3 ±0.4%
C-LSTM	25 Secs	94.2%	85.0%	**89.3** ±0.5%
Anom. Trans.	5 Secs	91.0%	87.7%	**89.3** ±0.4%
Anom. Trans.	15 Secs	93.0%	94.2%	**93.6** ±0.8%
Anom. Trans.	25 Secs	88.4%	90.0%	89.2 ±0.3%

Table 3. Key model results: Kamps, Random Forest, C-LSTM, Anom. Trans.
Source: Chadalapaka Viswanath et. al., 2022

Their research findings indicate that predictions made using the 5-second chunked dataset are considerably less accurate than those made using the 15 and 25-second chunked datasets, implying that anomaly prediction using smaller chunk sizes is a more difficult problem in general. These results confirm prior research findings as well. Overall, this study highlights the potential of deep learning methods to revolutionize the detection of fraud in the rapidly evolving world of cryptocurrency.

Unfortunately, that’s the first and the last article about P&Ds detection via Deep Learning approach that was found. That’s why further research must be conducted in order to detect P&D schemes and, hence, save usual investors from such scam.

4. Anomaly Detection Models

Anomaly detection is a process of identifying data points, patterns, or events that deviate significantly from the expected behavior or norm. It can be used for various applications, including fraud detection, system monitoring, and predictive maintenance. There are various approaches to anomaly detection, including statistical methods, machine learning algorithms, and rule-based systems. Statistical methods rely on analyzing the distribution of data to identify outliers. Machine learning algorithms use supervised or unsupervised techniques to train a model to identify anomalies. Rule-based systems use pre-defined rules or thresholds to flag anomalies. Anomaly detection has become increasingly important in today's data-driven world, as it can help organizations identify and address potential risks or problems before they cause significant damage. However, it is important to note that anomaly detection is not a perfect science and can sometimes produce false positives or miss true anomalies. Therefore, it is crucial to continuously evaluate and refine anomaly detection methods to improve accuracy and effectiveness.

Anomalies or outliers are data points that deviate from the usual patterns in a dataset as it was already mentioned. Detecting these anomalies can be achieved through supervised or unsupervised techniques.

In supervised techniques, the system learns from a set of normal training data with already collected target variable, while unsupervised techniques assume that anomalies are rare and require the analyst to determine the parameters for identifying them, so, there is no labelled target variable. It can be challenging to obtain an adequate training set for supervised anomaly detection, but we will do it in our paper using different Python[25] libraries and brute force approach, we will talk about it in the next chapter.

P&D scheme is considered an anomaly because it goes against the principles of fair market practice and can swindle investors out of their hard-earned money. Financial analysts and regulators use various methods to detect and prevent Pump and Dump Schemes, such as monitoring trading volumes, analyzing social media activity, and conducting investigations. However, these methods are not foolproof, and some schemes can still go undetected. That’s why it is important to deep into anomaly detection methodology and explore the most popular and effective methods that will help to detect and prevent such hazardous activities on crypto markets.

[25] Python webpage. Available: https://www.python.org.

4.1 Types of Anomalies

There are different types of anomalies, which have been grouped into 3 major categories by Chandola et al., 2009[26]: collective anomalies, contextual anomalies, and point anomalies.

Point anomalies are individual data points that deviate from the usual trend of the dataset, such as an exceptionally large purchase compared to a person's typical spending pattern, that’s actually the case of P&D scheme. In contrast, collective anomalies occur when a group of data points occur together in time or sequence, indicating abnormal behavior. An example of this would be a series of unusually low readings on an electrocardiogram indicating a potential problem rather than a single reading alone. Finally, contextual anomalies represent data points that would only be considered unusual in specific situations or conditions, such as a warm temperature in winter, which would be unexpected but not abnormal in summer.

4.2 Machine Learning Models to Detect Anomaly

Anomaly detection is a crucial task in machine learning applications across various domains such as cybersecurity, finance, health monitoring, and more. Anomaly detection models are developed to identify unexpected patterns or outliers in a given dataset. These models can use supervised or unsupervised learning methods to identify anomalies as it was already discussed.

We will discuss the main machine learning models used for anomaly detection, let’s start with unsupervised methods:

1. Isolation Forest
2. One-Class SVM
3. Local Outlier Factor (LOF)

Isolation Forest[27] is an unsupervised algorithm that is based on trees. Its primary function is to identify anomalies in a dataset by isolating them from a group of normal observations. This model is efficient, fast, and can handle large datasets. It works by creating a random binary tree, in which nodes are partitioned randomly in a way that each partition isolates the data points into different sub-spaces. By repeating this process for a number of times, anomalies are more likely to have a shorter path length in the trees, making them easier

[26] «Anomaly detection: A survey», Chandola, V., Banerjee, A., & Kumar, V, 2009, ACM Computing Surveys (CSUR), 41(3), 15.

[27] «Random Decision Forests», Ho, Tin Kam, 1995, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Pp. 278-282

to detect as they are isolated from the rest of the data points. It does not rely on any assumptions about the distribution of the data, and it can handle both continuous and categorical features. However, like any other algorithm, Isolation Forest has its limitations. It may not perform well with highly overlapping or clustered data, and it requires tuning of parameters such as the number of trees and the maximum depth of the trees. It is also important to note that Isolation Forest, like any other algorithm, is not perfect and can sometimes produce false positives or miss true anomalies. Therefore, it is crucial to combine it with other anomaly detection techniques and domain expertise when using it in practice.

One-Class SVM is based on support vector machines (SVMs) and is also an unsupervised algorithm. It is a binary classification model that is used to detect anomalies. The algorithm uses a single-class label, which is the normal data, and then identifies data points that fall outside of that class label as anomalies.

Local Outlier Factor (LOF) LOF is an unsupervised algorithm that assesses the local density of the dataset. It identifies data points that have a lower density than their surrounding data points as anomalies.

So, the models such as Isolation Forest, One-Class SVM, LOF can effectively identify anomalies in datasets for various domains. Selecting an appropriate model or combination of models depends on the specific task, data, and application requirements. Nevertheless, in this paper we will focus on supervised research. We will collect pre-labelled data.

That’s why we must delve into the supervised machine learning methods, here are some of the main well-known methods that can help us to detect pump and dump schemes:

1. Support Vector Machines (SVM)
2. Random Forests
3. Logistic Regression

SVM is a powerful algorithm used for classification and regression analysis. SVMs work by finding the optimal separating boundary between two classes of data points. Anomaly detection using SVMs is done by identifying points that fall on the wrong side of the boundary.

Random Forests is an ensemble learning algorithm that combines multiple decision trees to produce a more accurate result. In the Random Forest algorithm, each decision tree is built using a random subset of features from the original dataset, which helps to reduce the chances of overfitting and improve the overall accuracy of the model. Additionally, the decision trees are built independently of each other, which makes the algorithm highly scalable and suitable for large datasets. The process of selecting the best feature to split the data is repeated recursively until the data is split into smaller and smaller subsets, or until a stopping criterion is met. Once all the trees in the forest have been built, the algorithm combines the predictions from each tree to produce a final result. In classification tasks, the final prediction is based on the majority vote of all the trees, while in regression tasks, it is based on the average of all the tree predictions. Random Forest models are known for their high accuracy, robustness, and ability to handle noisy data. They are widely used in various fields such as finance, medicine, and image recognition. However, they can be computationally expensive and require careful tuning of hyperparameters to achieve optimal performance. In anomaly detection, random forests are used to identify patterns in the data that are uncommon, thus indicating an anomaly.

Logistic regression is a commonly used statistical method for anomaly detection. It is a type of supervised learning algorithm that can be trained using a set of labeled examples of normal and anomalous data. Once trained, the logistic regression model can predict the probability whether a new observation is normal or anomalous based on the inputs given to it. If the predicted probability exceeds a certain threshold, the observation can be flagged as potentially anomalous and further investigation can be conducted.

It's important to note that logistic regression, like any machine learning methods, is not foolproof and can produce false positives or miss true anomalies. Therefore, it is crucial to carefully select and validate the input features, train the model on a diverse and representative dataset, and continually monitor and refine the detection algorithm over time.

4.3 Deep Learning Models to Detect Anomaly

Deep learning models have revolutionized anomaly detection in recent years. That’s because they outperform classical machine learning methods in any metrics. These models use artificial neural networks with multiple layers to learn from vast amounts of data and identify patterns and anomalies.

Here are some of the popular deep learning models used for anomaly detection:

1. Autoencoders
2. Long Short-Term Memory (LSTM)
3. Convolutional LSTM (C-LSTM)

Autoencoders are neural networks that aim to reconstruct inputs, unsupervised learning algorithm. They learn to represent the input data through an encoding process and then reconstruct it using a decoding process. During training, an autoencoder learns to preserve as much information as possible while simultaneously eliminating noise and outliers. This ability to encode and decode data makes autoencoders an excellent tool for identifying anomalies.

LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that is used for processing sequential data, such as time series data, natural language processing, and speech recognition. LSTM networks are designed to overcome the limitations of traditional RNNs, which struggle to capture long-term dependencies in sequences. LSTM networks use a memory cell, which is a unit that can store information over time and selectively forget or remember information based on the input data. This memory cell is controlled by three gates: the input gate, the output gate, and the forget gate. These gates control the flow of information into and out of the memory cell, allowing the LSTM network to selectively retain or discard information from the input sequence. The input gate controls how much new information is added to the memory cell, while the forget gate determines how much information is discarded from the memory cell. The output gate controls how much information is extracted from the memory cell and used as the output of the LSTM network. The learning process in LSTM networks involves training the network to adjust the weights and biases of the gates to optimize the performance of the network on a given task. This is typically done using backpropagation through time (BPTT), which is a variant of the backpropagation algorithm used for training traditional neural networks. LSTM networks have been shown to be effective for a wide range of tasks, including speech recognition, language translation, and image captioning. They are particularly useful for tasks that involve processing long sequences of data, such as natural language processing, as they are able to capture long-term dependencies in the data. However, they can be computationally expensive and require a large amount of training data to achieve good performance.

C-LSTM (Convolutional LSTM) is a neural network architecture that combines convolutional layers and LSTM layers for processing sequential data such as time series or text. C-LSTM includes a convolutional layer that extracts features from sequential data and an LSTM layer that processes these features while retaining information about previous states. C-LSTM has shown good results in time series forecasting tasks. However, C-LSTM has a more complex architecture than simple LSTM networks, which can lead to longer training times and more complex parameter tuning. Here is a visual form of C-LSTM model:

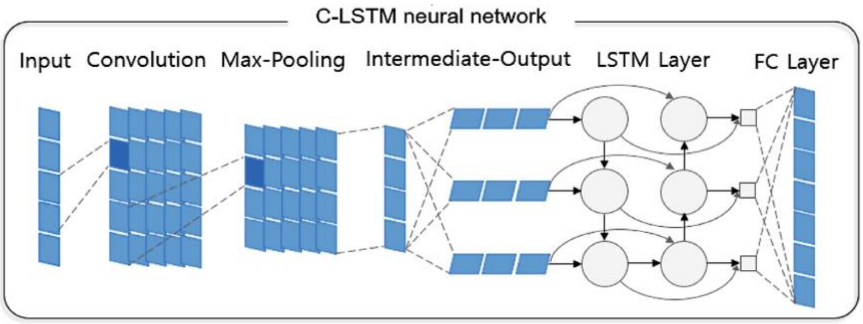


Figure 8. C-LSTM model architecture
Source: Tae-Young Kim et al., 2018[28]

Besides, there are also the most recent articles about deep learning techniques that can detect and predict anomalies that outperform state-of-the-art deep learning model such as LSTM that was already mentioned. For example, TranAD[29]. TranAD is a novel deep learning model for detecting and diagnosing anomalies in data. It is based on transformer networks and uses attention-based sequence encoders to rapidly perform inference while taking into account broader temporal trends in the data. To ensure robust feature extraction, TranAD utilizes focus score-based self-conditioning and adversarial training for greater stability. Moreover, TranAD incorporates model-agnostic meta learning (MAML) to train the model with limited data. Here is an architecture of the model:

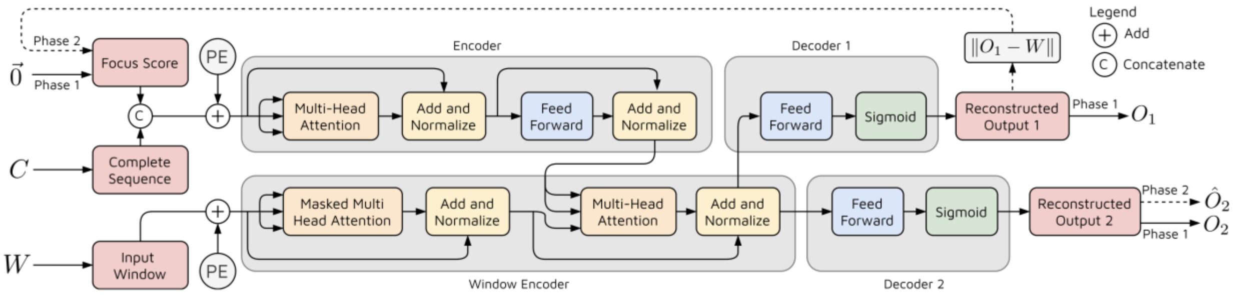


Figure 9. TranAD model architecture
Source: Shreshth Tuli at el., 2022

Extensive experiments on six publicly available datasets indicate that TranAD outperforms state-of-the-art baseline methods in terms of detection and diagnosis performance. Specifically, TranAD can increase F1 scores by up to 17% and reduce training times by up to 99% compared to the baselines such as LSTM, making it a highly efficient and effective tool for anomaly detection.

[28] «Web traffic anomaly detection using c-lstm neural networks», Tae-Young Kim and Sung-Bae Cho. 2018, Expert Systems with Applications, 106:66–76.
[29] «TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data», Shreshth Tuli, Giuliano Casale, Nicholas R. Jennings, 2022, Proceedings of the VLDB Endowment

There are lots of deep learning algorithm can be found in the article that is called «Deep Learning for Anomaly Detection: A Review»[30]. This article provides a review of deep anomaly detection research, detailing a comprehensive classification system that covers progress in three high-level categories and eleven fine-grained categories of the methods.

4.4 Evaluation metrics

Accuracy is not enough, that's the key idea. In anomaly detection applications the distribution between normal and abnormal classes can often be highly imbalanced. This challenge is commonly referred to as the class imbalance problem. Learning from such skewed data can result in a model that is accurate when classifying examples within the normal class, but may perform poorly when classifying anomalous examples.

For instance, let's consider a dataset consisting of 1,000 images of luggage passing through a security checkpoint. Out of these images, 950 are of normal luggage, and 50 are abnormal. A classification model that always predicts an image as normal can achieve high overall accuracy (95%) for this dataset, but its accuracy rate for classifying abnormal data would be 0%. Moreover, such a model may misclassify normal examples as anomalous (false positives, FP) or anomalous examples as normal (false negatives, FN). As we account for both these types of errors, we can conclude that the traditional accuracy metric, which divides the total number of correct classifications by the total classifications, is insufficient to evaluate the effectiveness of an anomaly detection model.

Two important metrics have been introduced to improve measuring model skill: precision and recall. Precision is defined as TP divided by TP + FP, while recall is TP divided by TP + FN. Depending on the application, it may be preferable to optimize for either precision or recall. Optimizing for precision is useful when failure cost is low or to decrease human workload. Optimizing for high recall is appropriate when cost of false negatives is very high, such as in airport security. The way the threshold is set can reflect the precision and recall preferences for each specific use case.

Nevertheless, the most important metric for us is a F1-score. F1 score is a metric used to evaluate the performance of a binary classification model. It is the harmonic mean of Precision and Recall. It is useful in situations where both Precision and Recall are important and need to be considered together. A high F1 score indicates that the model has a good balance

[30] «Deep Learning for Anomaly Detection: A Review», Guansong Pang, Chunhua Shen, Longbing Cao, Anton Van Den Hengel, 2021, ACM Computing Surveys, Volume 54, Issue 2, Article No.: 38, 1–38

between Precision and Recall, while a low F1 score indicates that the model may be biased towards one of the metrics. We will mainly compare the results according to F1-score.

In anomaly detection, ROC-AUC (Receiver Operating Characteristic Area Under the Curve) is also a commonly used metric to evaluate the performance of a classification model. ROC-AUC measures the ability of the model to distinguish between normal and anomalous data points. ROC-AUC is a suitable metric for anomaly detection when the dataset is imbalanced, meaning that the normal data points significantly outnumber the anomalous ones. ROC-AUC takes into account both sensitivity (true positive rate) and specificity (true negative rate) of the model. A higher ROC-AUC score indicates that the model has better discrimination ability, i.e., it can distinguish between normal and anomalous data points more accurately. Therefore, using ROC-AUC in anomaly detection is a valid and appropriate approach for evaluating the performance of a classification model. That's why we also will calculate ROC-AUC metrics for the main models.

5. Pump and Dump Detection

5.1 Idea

As we have already understood, usual investors often fall victim to P&D scheme. When they can see on exchange or some additional website with the data of different type coins such as CoinMarketCap[31] a cryptocurrency's price rise, they may see it as a promising investment opportunity. However, that's not the case during a P&D, as the rise lacks economic basis and is merely a product of market manipulation. To safeguard investors, it is crucial to determine whether a P&D scheme can be detected and how good according to appropriate metric it can be done. This is the objective of this Chapter.

To gain a better understanding of how pump and dumps can be detected, it is essential to have some basic knowledge. Let's start with the order book for the cryptocurrency. The order book displays the total quantity of pending buy and sell orders for coins at each price above and below the market price. Asks are sorted from lowest to highest price, while bids are sorted from highest to lowest. The quickest way to purchase on the market is through a buy market order, so, it is the best way for investors in these Telegram chats to execute immediately their orders. There are some other types of orders such as limit order to buy that buy at specific price. However, it's impossible to use limit orders or the similar ones while there is no precise data when the Pump will be organized and what coin will be pumped, probably it can be used by admins or VIP members. Market order usually consist of the data about volume, price, side and timestamp. Collecting all such available information, I will try to detect the P&D scheme start.

5.2 Data

Nowadays it is possible to find few datasets that will help to fulfill the goals of this research, these datasets were published by the authors of articles such as «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations»[32] and "Sequence-Based Target Coin Prediction for Cryptocurrency Pump-and-Dump"[33]. However, it is crucial to create a new and unique dataset because of two main reasons:

[31] https://coinmarketcap.com

[32] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

[33] "Sequence-Based Target Coin Prediction for Cryptocurrency Pump-and-Dump", Hu, Sihao and Zhang, Zhen and Lu, Shengliang and He, Bingsheng and Li, Zhao, 2022, arXiv preprint arXiv:2204.12929

1. The most appropriate dataset from the first articles used P&Ds that were collected from 2017 to 2020, it is too outdated, the crypto market has changed a lot in the last few years. Hence, it is interesting to check is it possible to detect pump and dump scheme that was organized not so long ago.
2. The dataset from the second article used unsupervised approach to collect the P&Ds occurrences. Thus, mistakes in datases can exist because of different errors and gaps in data: postpone of P&D, announcing the coin in a sneaky OCR-proof image, etc..

There are several steps to create needed data. First of all, I must collect some P&Ds. To do so, we can use the website that is called PumbOlymp[34], it provides an extensive database of numerous Telegram channels for pump-and-dump schemes for free that were collected by searching pump-related keywords such as “pump”, “coin”, “vip”, “announcement” and so on. It is must to notice that not all these channels are related to pump and dump scheme. Furthermore, some of them are organize pumps only on specific exchanges such as Kucoin, Hotbit, Yobit, etc.. For example, there is a Telegram chat that is called “Kucoin Crypto Pumps”[35], this chat obviously organizes pump and dump schemes only on Kucoin exchange. That’s why it takes some time to find appropriate ones.

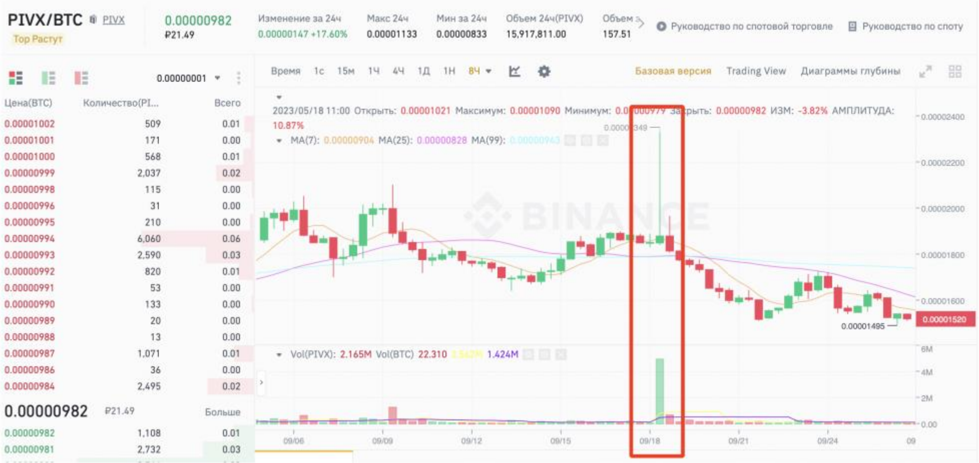
In this research we will focus only on P&Ds that were organized on Binance exchange. The reasons are that, firstly, only the Binance API gives an opportunity to collect historical data using ccxt library (Kucoin, Yobit, Hotbit and others are not available), secondly, it is best exchange according to volume trading per day and the overall rating, thus, the most challenging and interesting. Ccxt library – python library that provides an access to a big list of crypto exchanges and their data.

№	Channel	link	Number of subscribers by 14th May 2022
1	💎🚀 Hit Pump Angels 🚀💎	https://t.me/hitpumpangels	**14753**
2	Binance Crypto Pumps	https://t.me/Big_Pumps_Signals	**356618**
3	Big Pumps Binance	https://t.me/+dfOF0OmHl6YwMjQ9	**54996**
4	Binance Crypto Pump Signals🚨🚨🚨	https://t.me/BinancePumpSignalstrading	**13591**
5	[Official]Binance Pump 247 [Annoucement]	https://t.me/binancepump247	**196**

Table 4. Telegram chats
Source: Own analysis

[34] https://pumpolymp.com
[35] https://t.me/Kucoin_Crypto_Pumps

All in all, we joined 5 different Telegram chats (check Table 3) and collected 76 different P&Ds for the last 2 years that were organized on Binance exchange (check the Appendix). Here just an example of the pump with PIVX coin from the collected pumps:



PIVX/BTC PIVX
Top Растут

0.00000982
₽21.49

Изменение за 24ч	Макс 24ч	Мин за 24ч	Объем 24ч(PIVX)	Объем
0.00000147 +17.60%	0.00001133	0.00000833	15,917,811.00	157.51

Руководство по спотовой торговле
Руководство по споту

0.00000001

Цена(BTC)	Количество(PI...	Всего
0.00001002	509	0.01
0.00001001	171	0.00
0.00001000	568	0.01
0.00000999	2,037	0.02
0.00000998	115	0.00
0.00000996	31	0.00
0.00000995	210	0.00
0.00000994	6,060	0.06
0.00000993	2,590	0.03
0.00000992	820	0.01
0.00000991	53	0.00
0.00000990	133	0.00
0.00000989	20	0.00
0.00000988	13	0.00
0.00000987	1,071	0.01
0.00000986	36	0.00
0.00000984	2,495	0.02
0.00000982	₽21.49	Больше
0.00000982	1,108	0.01
0.00000981	2,732	0.03

2023/05/18 11:00 Открыть: 0.00001021 Максимум: 0.00001090 Минимум: 0.00000979 Закрыть: 0.00000982 ИЗМ: -3.82% АМПЛИТУДА: 10.87%
MA(7): 0.00000904 MA(25): 0.00000828 MA(99): 0.00000943
Vol(PIVX): 2.165M Vol(BTC) 22.310 1.424M

Figure 10. PIVX/BTC pump, 2021-09-18 15:00, announced growth: 100%
Source: Own analysis

This dataset proves once again some results from the literature review:

1. If the coin was once pumped, the probability of the future pump increases exponentially. There are more than 20 coins that were pumped more than one time, and 10 coins that were pumped more than twice,
2. Pump and dump schemes create significant market distortions, some coins were pumped with announced growth of more than 2000% in price. The average announced growth is equal to 163%.
3. Some pumps were organized by more than one Telegram chat. For example, approximately all pumps that were organized by “Big Pumps Binance” were also organized by “[Official]Binance Pump 247 [Annoucement]”, that’s why the only one unique pump was included in the final dataset from the second Telegram chat. Probably the second Telegram chat was created by admins of the first chat. The same type of cooperation by admins can be found within “Hit Pump Angels” and “Binance Crypto Pumps”.
4. Alt-coins are the main coins that used by admins of the Telegram chats for the P&Ds.

Once the dataset with pumps were created, using timestamps of pumps and ccxt library, all available trades for up to 2 days preceding and succeeding the pump were collected. The reason why we collected only 4 days for each pump is the following: not to include one more pump in dataset. According to collected data frame and literature review different groups can arrange one more pump on the same alt-coin after a few days, plus to it there is lots of groups that we have not researched because of different reasons such as language barrier, they can also organize the pump again on the same coin. Besides, we figure out that not all transaction can be downloaded using ccxt, there were huge gaps in data for some coins, it is probably because of some problems in ccxt library. That's why, only 43 of them left.

After collecting all available data for 43 pumps, we further preprocess it by aggregating the trades in 15-seconds chucks and introducing sliding window as it was done by Massimo La Morgia et al.[36]. According to the literature review it is one of the best chunk size for pump detection according to F1 metric. It is a must to notice that we analyzed only buy orders because of the main goal – to detect the start of the pump, and obviously it starts with buy orders.

Unfortunately, ccxt library and its access through Binance API do not give the information about the type of orders that were placed by the investors (Limit, Market, etc.), that's why we will introduce "rush orders". Usually, such types of orders execute in a single shot, so, to account for them, we can aggregate all trades that were executed in one time as one trade. Thus, if there were, for example, more than 10 trades at the same millisecond, then it is our 1 rush order.

Finally, here are the features that I used, they are approximately like the features from Massimo La Morgia et al., that's because we also want to check the consistency and comparability of the past results on the new dataset that was created:

1. ***Date, hour_sin, hour_cos, minute_sin, minute_cos:***: The timestamp of order and all related type features
2. ***Pump_index***: Number of P&D scheme,
3. ***Symbol***: The name of the coin
4. ***StdTrades***: Moving stardard deviation of the number of trades,
5. ***StdVolumes***: Moving standard deviation percent change of volume of trades,
6. ***AvgVolumes***: Moving average percent change of volume of trades,
7. ***StdRushOrders***: Moving standard deviation percent change of volume of rush orders,
8. ***AvgRushOrders***: Moving average of volume percent change of rush orders,
9. ***StdPrice***: Moving standard deviation of closing price,
10. ***AvgPrice***: Moving average percent change of closing price,
11. ***AvgPriceMax***: Moving average percent change of maximal price

In summary, we collected dataset with 203,652 rows for 43 pumps. The last step is to label target variable manually – "Was Pump or No". The reason of it is that during the analysis, we found that in some cases, P&Ds started before or after the organizer shared the signal within 15 seconds gap. Nevertheless, that's all about few cases (only 8 out of 43), hence, approximately all pumps were labeled according to collected timestamp. The final version of the dataset and Python code is released to my GitHub[37].

[36] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

[37] https://github.com/TimurPugoev

5.3 Methodology

Before we delve into the empirical research, we must set the hypothesis that should be checked:

1. **H1**: It is possible to detect the P&D scheme,
2. **H2**: Current deep learning models outperform basic machine learning models such as Random Forest in P&D start detection.

To detect the start of P&Ds with the collected dataset with several features and labeled target variable, we can use different types of supervised learning models as it was discussed in the previous chapter "Anomaly Detection Models".

Firstly, we will focus on the Classical Machine Learning techniques. According to the literature review and research of anomaly detection models, the Random Forest is the best model according to F1 metric in P&Ds detection. Even though Logistic Regression also provides significant results, this model performed worse, hence, we will not include it in the analysis. Besides, we must include SVM model because there is no article that used such model in P&Ds detection and, thus, it will be interesting to check what result can be performed on this type of data.

Secondly, we will consider Deep Learning models. Our analysis will start with C-LSTM neural network architecture that was already used on the Massimo La Morgia, 2020[38] dataset by Chadalapaka Viswanath et. al., 2022[39]. There are 2 main reasons why such model will be included in this research:

1. As it was already mentioned in the literature, such model outperforms classical models on the outdated dataset. Hence, it will be interesting to compare the past results with the results that were performed using the new data,
2. We will use it as the baseline in the deep learning models. Besides, we will have an appropriate opportunity to check one of the hypotheses: deep learning models outperform classical state-of-the-art machine learning models.

Furthermore, the different configuration of the C-LSTM model will be researched to check the sensitivity of results to parameters. Finally, here are the list of the models that will be used in this research:

[38] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

[39] «Crypto Pump and Dump Detection via Deep Learning Techniques», Chadalapaka Viswanath, Chang Kyle, Mahajan Gireesh, Vasil Anuj, 2022, arXiv

1. Random Forest,
2. SVM,
3. C-LSTM.

Let’s deep further into the methodology. Firstly, we will use the same methodology that was introduced by Massimo La Morgia, 2020 with some additional analysis. The authors did not split their dataset into train/test sets, they just performed cross-validation with k folds, where k is 5 or 10 according to the time scale of the chunk. For the 15-seconds chunks they used 5 folds cross-validation, the same we will apply in this research. They used Random Forest with the following parameters: 200 trees, maximum depth is equal to 4. We will also use other parameters of RF [number of trees: 50, 100, 150; maximum depth: 5] to check if it’s possible to outperform such parameters in terms of metrics: F1 score and a timeframe of the overall work of the model.

Secondly, we will apply the next methodology that was introduced by Chadalapaka Viswanath et. al., 2022. They divide the data on the train and test split using 80:20 ratio without shuffling taking into consideration the sequence X = X_1, X_2…X_N, where N is the number of pumps. Separating according to pumps is a must to ensure that model is not fed information from two pumps in one time. Then they exclude the X_i that has less than 100 rows for the training dataset. However, our final dataset consists only well-selected data, that’s why 43 out 43 pumps left. For the classical machine learning models, they just recomputed using the same 80:20 split, the same we will also do. For the deep leaning models, they created a unique approach that we will also apply. They proceed to preprocess the information for each pump by segmenting them into parts of size s through the implementation of a sliding window over the chunks of each pump X_i. Additionally, reflection padding was added to the beginning of each segment to increase its size by s-1, guaranteeing that the total number of windows remained consistent with the original count of chunks. These segments were subsequently utilized as inputs for the models, which predicted the probability of a pump taking place during the final chunk of the segment. Importantly, this preprocessing strategy permits us to safely shuffle and batch segments from all pumps without the risk of cross-contamination between different training data sets. Foremost, segmenting and predicting in this manner minimizes the possibility of our models making predictions based on future values. In other words, our models can only predict the occurrence of a pump based on presently available information. This safeguards the efficacy of our models, enabling them to be used in real-world, live anomaly detection scenarios on actual exchanges. We will also use undersampling to boost the performance of the model. However, for the validation no undersampling is performed to avoid bias in metrics.

Here is the architecture of the C-LSTM model that was used by Chadalapaka Viswanath et. al., 2022, the same we will apply for our dataset:

- 1 LSTM layer with an embedding dimension of 350,
- 1 set of convolutional/ReLU/pooling layers with a convolution kernel size of 3 with a stride of 1,
- A pooling kernel size of 2 with a stride of 1,
- 1 feedforward layer which directly projects the last hidden state of the LSTM to a dimension of 1.
- A sigmoid layer which constrains the output of our classifier between 0 and 1.

5.4 Results

Using the first methodology we have the following result of Random Forest model with different parameters:

Chunk size	Parameters	Precision	Recall	F1-score	Time spent
15s	▪ Trees: 50 ▪ Max depth: 4	89.2%	78.6%	83.5%	2.29 seconds
15s	▪ **Trees: 50** ▪ **Max depth: 5**	**89.2%**	**78.6%**	**83.5%**	**2.28 seconds**
15s	▪ Trees: 100 ▪ Max depth: 4	86.8%	78.6%	82.5%	3.75 seconds
15s	▪ Trees: 100 ▪ Max depth: 5	84.6%	78.6%	81.5%	4.09 seconds
15s	▪ Trees: 150 ▪ Max depth: 4	84.6%	78.6%	81.5%	5.51 seconds
15s	▪ Trees: 150 ▪ Max depth: 5	84.6%	78.6%	81.5%	5.96 seconds
15s	▪ Trees: 200 ▪ Max depth: 4	84.6%	78.6%	81.5%	8.07 seconds
15s	▪ Trees: 200 ▪ Max depth: 5	86.8%	78.6%	82.5%	7.96 seconds
Massimo La Morgia et. al., 2020[40] results:					
15s	▪ Trees: 200 ▪ Max depth: 4	91.3%	84.4%	87.7%	5 seconds

Table 5. Random Forest (5 Folds), Results
Source: Own analysis

[40] «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)

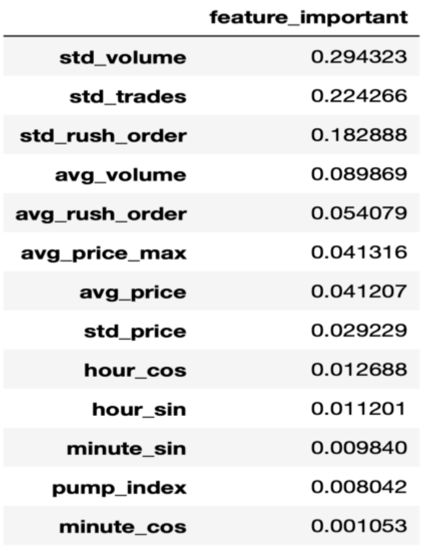
As we can see, F1-score, Precision and Recall for each type of parameters with our created dataset is less than the identical metrics that was reached by Massimo La Morgia et. al. with their outdated dataset. That's actually reasonable because the crypto market grew a lot in the last 5 years [most of the pumps from their dataset were organized in 2017-2018], hence, it should be more complicated to detect P&Ds schemes. However, the difference is not so significant. Our best models in terms of F1-score is a Random Forest model with 50 trees and maximum depth that is equal to 5, other metrics are also the best over the other configurations. Anyway, now we can easily prove the first hypothesis H1, that's possible to detect pump and dump schemes.

Let's move further and check the second hypothesis H2. Using the second methodology and previous results, firstly, let's split the data for the baseline model – Random Forest with $trees = 50$ and $maximum\ depth = 5$. Thus, 33 first pumps will be used for the training set, other for the test set. The same split will be used for the SVM model. Table 5 presents the results of the analysis:

Model	Precision	Recall	F1-score	ROC-AUC
SVM	100%	33.3%	50%	66.6%
Random Forest 50 trees, max depth 5	72.7%	88.8%	79.9%	94.4%

Table 6. Random Forest, SVM Results
Source: Own analysis

We see that according to F1-score Random Forest outperform SVM significantly, the same with ROC-AUC and Recall. In Figure 11 we list feature importance based on Gini impurity:



	feature_important
std_volume	0.294323
std_trades	0.224266
std_rush_order	0.182888
avg_volume	0.089869
avg_rush_order	0.054079
avg_price_max	0.041316
avg_price	0.041207
std_price	0.029229
hour_cos	0.012688
hour_sin	0.011201
minute_sin	0.009840
pump_index	0.008042
minute_cos	0.001053

Figure 11. Feature importance, Random Forest
Source: Own analysis

The most 3 important features are StdVolume, StdTrades, StdRushOrders. That's coincide with the results from the literature review, where it was mentioned that one of the most important changes that caused pump is the change in the volume [Pre-pump volume VS Pumped volume, check Figure 5]. It also proves the significance of the created "rush" order feature.

Well, in the next step we have to evaluate the C-LSTM model and reject or approve the second hypothesis. We used batch size of 600, undersampling proportion of 0.05, precision-recall threshold of 0.4 and number of epochs that is equal to 50. Table 5 present the best result of our C-LSTM model that was used on our new created data and C-LSTM model that was used by Chadalapaka Viswanath et. al. on outdated data:

Dataset	Model	Precision	Recall	F1-score
New	C-LSTM	80%	**88.9%**	84.2%
Old	C-LSTM	**94.2%**	84.9%	**89.3%**

Table 7. C-LSTM results
Source: Own analysis

We have the same tendency again, F1-score for the new created dataset decreased comparing with the old one dataset. However, if we compare the results of C-LSTM and Random Forest we can easily approve the second hypothesis H2 that deep learning models outperform classical machine learning models because there is an increase in approximately 5% of F1-score.

It is also interesting to check possible results of C-LSTM model with other parameters, one of the most important is a precision-recall threshold. We will, firstly, apply such configurations of this parameter: 0.2, 0.8. Secondly, we will introduce "pick_threshold" function that will choose the parameter by itself according to F-1 score. The Table 5 presents the most interesting results. We see that using the 0.2 as a threshold decrease overall results to the similar one from the Random Forest. The threshold that was found by the function in the 30-th epoch learning equals to 0.49, the results are the like the results with 0.4. Another interesting result we can see on the last row of the table, if we have high threshold, then our score can decrease. There are also configurations as batch size, undersampling proportion, number of epochs and even the initial configuration of C-LSTM architecture to change, we leave it to the further research papers.

Dataset	Model	Precision	Recall	F1-score	ROC-AUC
New	C-LSTM Prec-Rec threshold = 0.2	72.7%	88.8%	79.9%	94.4%
New	C-LSTM Prec-Rec threshold = 0.8	80%	88.9%	84.2%	94.44%
New	C-LSTM Prec-Rec threshold, found by the function = 0.497	80%	88.9%	84.2%	94.44%
New	C-LSTM Prec-Rec threshold, found by the function = 0.905	75%	33.3%	46.1%	66.6%

Table 8. C-LSTM with different parameters
Source: Own analysis

6. Conclusion

Based on the results obtained using Random Forest, SVM, and C-LSTM models, it can be concluded that it is possible to detect cryptocurrency pump-and-dump schemes, even in conditions when the crypto market has grown significantly. The neural network models, particularly C-LSTM, outperformed the other classical state-of-the-art machine learning models in terms of F1 score, demonstrating their potential in detecting and modeling such schemes. The study highlights the importance of using machine learning techniques in detecting fraudulent activities in the cryptocurrency market. Furthermore, the results highlight the importance of using multiple anomaly detection techniques and their parameters, such as C-LSTM precision-recall threshold, to, firstly, enhance the quality of detection according to some appropriate metric and, secondly, check the consistency of results.

Within the research, unique dataset was also created. The dataset consisted of manually labeled transaction data, providing a valuable resource for future research in this field. The creation of this dataset involved a significant amount of effort and attention to detail, highlighting the importance of high-quality data in training and evaluating machine learning models. The dataset was carefully curated to ensure that it represented a diverse range of pump-and-dump schemes according to the announced growth, thereby increasing the generalizability and validity of the study's findings. The utilization of this unique dataset also highlights the potential for future research to gather and analyze larger, more diverse data to further improve the performance and robustness of machine learning models in detecting and forecasting fraudulent activities in the cryptocurrency market.

Overall, the study's use of a unique dataset in conjunction with advanced machine learning techniques provides a valuable contribution to the field of cryptocurrency market analysis and has significant implications for the development of more secure and reliable cryptocurrency markets in the future. The findings of this study can be useful for investors, regulators, and law enforcement agencies in identifying and preventing pump-and-dump schemes in the cryptocurrency market.

Further researchers can focus on exploring other machine learning techniques such as TranAD model and incorporating additional data sources in order to detect and then prevent such hazardous activities.

5. Literature

1. «The Anatomy of a Cryptocurrency Pump-and-Dump Scheme», J Xu, B Livshits, 2019, USENIX Security Symposium
2. «Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations», Massimo La Morgia, Alessandro Mei, Francesco Sassi, Julinda Stefa, 2020 29th International Conference on Computer Communications and Networks (ICCCN)
3. «TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data», Shreshth Tuli, Giuliano Casale, Nicholas R. Jennings, 2022, Proceedings of the VLDB Endowment
4. «To the moon: defining and detecting cryptocurrency pump-and-dumps», J. Kamps, B. Kleinberg, 2018, Crime Science
5. «Stock manipulation and its effects: pump and dump versus stabilization», Yu Chuan Huang, Yao Jen Cheng, 2015, Review of Quantitative Finance and Accounting, volume 44, pages 791–815
6. «Unchecked intermediaries: Price manipulation in an emerging stock market», Asim Ijaz Khwaja, Atif Mian, 2005, Journal of Financial Economics, volume 78, pages 203-241
7. «Cryptocurrency Pumping Predictions: A Novel Approach to Identifying Pump And Dump Schemes», Cameron Ramos, Noah Golub, 2017, Stanford students, Final project
8. «Crypto Pump and Dump Detection via Deep Learning Techniques», Chadalapaka Viswanath, Chang Kyle, Mahajan Gireesh, Vasil Anuj, 2022, arXiv
9. «Price manipulation in the bitcoin ecosystem», N. Gandal, J. Hamrick, T. Moore, T. Oberman, 2018, Journal of Monetary Economics, vol. 95, pp. 86–96.
10. «An experimental study of cryptocurrency market dynamics» P. M. Krafft, N. Della Penna, A. S. Pentland, in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 2018, p. 605.
11. «The Economics of Cryptocurrency Pump and Dump Schemes», Feder, Amir & Gandal, Neil & Hamrick, JT & Moore, Tyler & Mukherjee, Arghya & Rouhi, Farhang & Vasek, Marie, 2018, CEPR Discussion Papers 13404, C.E.P.R. Discussion Papers
12. «A New Wolf in Town? Pump-and-Dump Manipulation in Cryptocurrency Markets», Dhawan, Anirudh and Putnins, Talis J, 2021, Review of Finance, Forthcoming

13. T. LLC. (2018) Telegram - a new era of messaging. Available: https://telegram.org/
14. Python webpage. Available: https://www.python.org.
15. Scikit-Learn webpage. Available: https://scikit-learn.org/stable/
16. «Profitability of cryptocurrency Pump and Dump schemes» Tsuchiya, T., 2021, Digit Finance 3, 149–167
17. «Anomaly detection: A survey», Chandola, V., Banerjee, A., & Kumar, V, 2009, ACM Computing Surveys (CSUR), 41(3), 15.
18. «Random Decision Forests», Ho, Tin Kam, 1995, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 278-282.
19. «Deep Learning for Anomaly Detection: A Review», Guansong Pang, Chunhua Shen, Longbing Cao, Anton Van Den Hengel, 2021, ACM Computing Surveys, Volume 54, Issue 2, Article No.: 38, 1–38
20. CoinMarketCap - The World's Most Popular Crypto Data Distribution Site. Available: https://coinmarketcap.com
21. «Sequence-Based Target Coin Prediction for Cryptocurrency Pump-and-Dump», Hu, Sihao and Zhang, Zhen and Lu, Shengliang and He, Bingsheng and Li, Zhao, 2022, arXiv preprint arXiv:2204.12929
22. «Web traffic anomaly detection using c-lstm neural networks», Tae-Young Kim and Sung-Bae Cho. 2018, Expert Systems with Applications, 106:66–76.

7. Appendix

Number_of_P&Ds	channel	coin	exchange	pair	timestamp	announced_growth
1	💎🚀 Hit Pump Angels 🚀💎	PIVX	binance	BTC	2021-03-28 17:00	295,36%
2	💎🚀 Hit Pump Angels 🚀💎	PPT	binance	BTC	2021-03-07 17:00	457,85%
3	💎🚀 Hit Pump Angels 🚀💎	MTH	binance	BTC	2021-03-01 17:00	51,11%
4	💎🚀 Hit Pump Angels 🚀💎	BRD	binance	BTC	2021-02-24 17:00	86,36%
5	💎🚀 Hit Pump Angels 🚀💎	RDN	binance	BTC	2021-02-22 17:00	82,46%
6	💎🚀 Hit Pump Angels 🚀💎	NXS	binance	BTC	2021-02-21 17:00	399,75%
7	💎🚀 Hit Pump Angels 🚀💎	PNT	binance	BTC	2021-01-31 21:00	1255,45%
8	💎🚀 Hit Pump Angels 🚀💎	EVX	binance	BTC	2021-01-09 21:00	350,80%
9	💎🚀 Hit Pump Angels 🚀💎	CTXC	binance	BTC	2021-01-07 15:59	100,46%
10	💎🚀 Hit Pump Angels 🚀💎	MDA	binance	BTC	2021-01-03 17:00	56,64%
11	💎🚀 Hit Pump Angels 🚀💎	NEBL	binance	BTC	2021-04-02 21:00	1256,58%
12	💎🚀 Hit Pump Angels 🚀💎	NXS	binance	BTC	2021-03-03 16:00	44,42%
13	💎🚀 Hit Pump Angels 🚀💎	DLT	binance	BTC	2020-12-26 21:00	602,66%
14	💎🚀 Hit Pump Angels 🚀💎	MTH	binance	BTC	2021-06-06 17:00	30,00%
15	💎🚀 Hit Pump Angels 🚀💎	RCN	binance	BTC	2021-01-07 21:00	197,54%
16	💎🚀 Hit Pump Angels 🚀💎	NAS	binance	BTC	2020-05-30 16:00	93,90%
17	💎🚀 Hit Pump Angels 🚀💎	VIA	binance	BTC	2020-05-29 15:59	214%
18	💎🚀 Hit Pump Angels 🚀💎	ONG	binance	BTC	2020-05-26 16:00	111,34%
19	💎🚀 Hit Pump Angels 🚀💎	ARN	binance	BTC	2020-04-07 18:00	29,52%
20	💎🚀 Hit Pump Angels 🚀💎	GRS	binance	BTC	2020-04-02 16:00	48,08%
21	💎🚀 Hit Pump Angels 🚀💎	EDO	binance	BTC	2020-03-26 15:59	49,02%
22	💎🚀 Hit Pump Angels 🚀💎	POA	binance	BTC	2020-03-12 19:00	56,25%
23	💎🚀 Hit Pump Angels 🚀💎	NXS	binance	BTC	2020-03-03 16:00	44,42%
24	💎🚀 Hit Pump Angels 🚀💎	RDN	binance	BTC	2020-02-01 13:00	30,56%

25	💎🚀 Hit Pump Angels 🚀💎	NULS	binance	BTC	2020-01-30 17:00	9,32%
26	💎🚀 Hit Pump Angels 🚀💎	SKY	binance	BTC	2021-02-03 21:00	2482,15%
27	💎🚀 Hit Pump Angels 🚀💎	FIO	binance	BTC	2021-02-04 17:00	62,10%
28	💎🚀 Hit Pump Angels 🚀💎	CTXC	binance	BTC	2021-03-18 17:00	54,75%
29	💎🚀 Hit Pump Angels 🚀💎	IDEX	binance	BTC	2020-12-31 17:00	132,75%
30	💎🚀 Hit Pump Angels 🚀💎	RCN	binance	BTC	2021-01-10 17:00	39,20%
31	💎🚀 Hit Pump Angels 🚀💎	ONG	binance	BTC	2021-01-11 16:00	52,60%
32	💎🚀 Hit Pump Angels 🚀💎	IDEX	binance	BTC	2021-01-17 17:00	30,43%
33	💎🚀 Hit Pump Angels 🚀💎	GVT	binance	BTC	2021-01-23 21:00	123,49%
34	💎🚀 Hit Pump Angels 🚀💎	VIB	binance	BTC	2021-02-05 21:00	561,53%
35	💎🚀 Hit Pump Angels 🚀💎	NEBL	binance	BTC	2021-02-13 21:00	128,78%
36	💎🚀 Hit Pump Angels 🚀💎	GVT	binance	BTC	2021-02-20 17:00	82,22%
37	💎🚀 Hit Pump Angels 🚀💎	ONG	binance	BTC	2021-04-12 17:00	16,60%
38	💎🚀 Hit Pump Angels 🚀💎	VIB	binance	BTC	2021-09-05 17:00	104,28%
39	💎🚀 Hit Pump Angels 🚀💎	OAX	binance	BTC	2020-10-18 18:00	43,87%
40	💎🚀 Hit Pump Angels 🚀💎	QSP	binance	BTC	2020-09-10 18:00	44,02%
41	💎🚀 Hit Pump Angels 🚀💎	QLC	binance	BTC	2020-09-06 16:00	60,25%
42	Binance Crypto Pumps	PHB	binance	BTC	2021-11-28 17:00	67%
43	Binance Crypto Pumps	APPC	binance	BTC	2021-12-19 15:00	45%
44	Binance Crypto Pumps	NEBL	binance	BTC	2022-01-02 17:01	70%
45	Binance Crypto Pumps	SNM	binance	BTC	2021-11-20 17:01	57,14%
46	Binance Crypto Pumps	NEBL	binance	BTC	2020-11-02 18:00	34,58%
47	Binance Crypto Pumps	PPT	binance	BTC	2020-06-08 16:06	54,05%
48	Binance Crypto Pumps	VIA	binance	BTC	2021-04-11 17:00	159,48%
49	Binance Crypto Pumps	DLT	binance	BTC	2021-05-16 17:00	25%
50	Binance Crypto Pumps	OAX	binance	BTC	2021-05-30 17:00	82,64%
51	Binance Crypto Pumps	WABI	binance	BTC	2021-06-20 17:00	108%
52	Binance Crypto Pumps	MTH	binance	BTC	2021-07-27 17:00	33,42%
53	Binance Crypto Pumps	NAS	binance	BTC	2021-08-22 17:00	157,68%
54	Binance Crypto Pumps	BRD	binance	BTC	2021-08-29 17:00	50%
55	Binance Crypto Pumps	FXS	binance	BTC	2021-09-19 17:00	100%
56	Binance Crypto Pumps	GVT	binance	BTC	2020-04-15 15:57	33%

57	Binance Crypto Pumps	BNT	binance	BTC	2020-04-24 15:59	17,79%
58	Binance Crypto Pumps	APPC	binance	BTC	2020-11-18 18:00	24,19%
59	Binance Crypto Pumps	MDA	binance	BTC	2020-11-27 21:00	21,67%
60	Binance Crypto Pumps	NAS	binance	BTC	2021-01-14 21:00	170,34%
61	Binance Crypto Pumps	DLT	binance	BTC	2021-01-28 19:00	30,40%
62	Binance Crypto Pumps	IDEX	binance	BTC	2021-04-25 17:00	40%
63	Binance Crypto Pumps	STEEM	binance	BTC	2021-01-18 17:00	16%
64	Binance Crypto Pumps	WPR	binance	BTC	2021-05-09 17:00	35%
65	Binance Crypto Pumps	FIO	binance	BTC	2021-06-13 17:01	36,13%
66	Big Pumps Binance	NEBL	binance	BTC	2021-08-24 15:00	93%
67	Big Pumps Binance	ATM	binance	BTC	2021-09-10 16:00	300%
68	Big Pumps Binance	PIVX	binance	BTC	2021-09-18 15:00	100%
69	Big Pumps Binance	WABI	binance	BTC	2021-10-14 15:00	85%
70	Big Pumps Binance	BRD	binance	BTC	2021-10-31 15:00	69%
71	Big Pumps Binance	NAS	binance	BTC	2021-11-14 15:00	75%
72	Big Pumps Binance	MDA	binance	BTC	2021-11-23 15:00	57%
73	Big Pumps Binance	SNM	binance	BTC	2022-01-27 15:02	17%
74	Big Pumps Binance	NAS	binance	BTC	2022-02-01 15:00	30%
75	Binance Crypto Pump Signals🚨🚨🚨	EVX	binance	BTC	2021-10-24 17:00	83,68%
76	[Official]Binance Pump 247 [Annoucement]	DREP	binance	USDT	2021-07-25 17:00	52,50%