future internet

MDPI

Review

A Survey on Pump and Dump Detection in the Cryptocurrency Market Using Machine Learning

Mohammad Javad Rajaei * and Qusay H. Mahmoud

Department of Electrical, Computer, and Software Engineering, Ontario Tech University, Oshawa, ON L1G 0C5, Canada; qusay.mahmoud@ontariotechu.ca
* Correspondence: mohammadjavad.rajaei@ontariotechu.net

Abstract: The popularity of cryptocurrencies has skyrocketed in recent years, with blockchain technologies enabling the development of new digital assets. However, along with their advantages, such as lower transaction costs, increased security, and transactional transparency, cryptocurrencies have also become susceptible to various forms of market manipulation. The pump and dump (P&D) scheme is of significant concern among these manipulation tactics. Despite the growing awareness of P&D activities in cryptocurrency markets, a comprehensive survey is needed to explore the detection methods. This paper aims to fill this gap by reviewing the literature on P&D detection in the cryptocurrency world. This survey provides valuable insights into detecting and classifying P&D schemes in the cryptocurrency market by analyzing the selected studies, including their definitions and the taxonomies of P&D schemes, the methodologies employed, their strengths and weaknesses, and the proposed solutions. Presented here are insights that can guide future research in this field and offer practical approaches to combating P&D manipulations in cryptocurrency trading.

Keywords: pump and dump detection; cryptocurrency; market manipulation detection; machine learning

Citation: Rajaei, M.J.; Mahmoud, Q.H. A Survey on Pump and Dump Detection in the Cryptocurrency Market Using Machine Learning. *Future Internet* **2023**, *15*, 267. https://doi.org/10.3390/fi15080267

Academic Editor: Sk. Md. Mizanur Rahman

Received: 17 July 2023
Revised: 30 July 2023
Accepted: 9 August 2023
Published: 11 August 2023

1. Introduction

The development of cryptocurrencies has altered financial markets and the way we see, interact with, and deal with digital assets [1] because of their decentralized organization and enhanced security. Compared to traditional fiat currencies, cryptocurrencies have a variety of benefits, including transparency, immutability, and borderless transactions. Several factors have contributed to the popularity of cryptocurrencies. To begin with, blockchain technology, which is the basis for most cryptocurrencies, has made recording and verifying transactions easier and more secure. Decentralized ledger systems eliminate intermediaries and centralized control, resulting in trust and transparency [2].

Despite its popularity and decentralized nature, cryptocurrency has also become a breeding ground for fraud. Since there are no centralized regulatory authorities and transactions are pseudonymous, cryptocurrency ecosystems are attractive targets for scammers. They exploit the lack of oversight and accountability to carry out these fraudulent practices. This is a threat to cryptocurrency investors and the overall stability of the market [3].

Fraudulent practices in cryptocurrency markets encompass various tactics. Spoofing entails the placement and subsequent cancellation of orders. Stop hunting targets stop-loss orders deliberately to influence price movements. Wash trading involves the execution of fictitious transactions to deceive market participants about liquidity and volume. Another notable scheme is P&D, where individuals or groups artificially inflate cryptocurrency prices before selling their holdings for personal gains, resulting in substantial losses for new investors.

Fake initial coin offers (ICOs) [4] have also been used to carry out fraudulent activities. A scammer creates ICO projects that look credible, promising high returns and cutting-edge technology. However, these projects often turn out to be fraudulent and do not have any

real development behind them, so when investors invest their money, they find out they have become victims of a fraud scheme.

It has also become increasingly prevalent to conduct phishing attacks against cryptocurrency users. Using various methods, hackers gain unauthorized access to cryptocurrency wallets and steal funds by tricking users into revealing their private keys or login credentials [5]. Social media analysis is of paramount importance in understanding and detecting fraudulent practices in cryptocurrency markets, as it can provide valuable insights into the manipulative strategies employed, enabling more effective measures against such schemes [6].

As mentioned earlier, in the cryptocurrency market, there is no regulation or oversight, which further exacerbates the problem of fraud. Although efforts are being made to establish regulatory frameworks and enhance security measures, cryptocurrencies' decentralized nature makes enforcing consistent standards and protecting investors difficult.

There are significant risks involved in these fraudulent activities for investors, and they undermine the integrity of the cryptocurrency market in general. The detection and prevention of such fraud is critical to protecting investors and maintaining the credibility of cryptocurrencies. Thus, the purpose of this survey is to explore the existing literature on fraud detection in cryptocurrencies. Our focus will be on detecting and analyzing P&D schemes, which have been identified as one of the most common and damaging forms of cryptocurrency fraud.

This study aims to provide valuable insights into the challenges and obstacles encountered in detecting and mitigating P&D schemes within the cryptocurrency ecosystem through a comprehensive analysis of the methodologies, techniques, and advancements presented in relevant research papers. In addition to recognizing the potential similarities and transferability of techniques, we also mention other notable research articles that have explored the detection of P&D schemes in various financial markets beyond cryptocurrencies. Eventually, the findings of this survey hold the potential to enhance the security and reliability of the cryptocurrency market, thereby instilling greater confidence among cryptocurrency investors.

This paper is organized into six sections. Section 1 introduces the topic and the motivation behind this study. Section 2 presents background and related works with a comparison of previous surveys on market manipulation and P&D detection and gives the theoretical background on cryptocurrency market manipulation and P&D schemes. Section 3 describes the methodology of the survey. Section 4 presents the various methodologies utilized and summarizes the results of relevant research. Section 5 presents and discusses the findings along with opportunities, challenges, and ideas for future research directions. Finally, concluding remarks are presented in Section 6.

2. Background and Related Work

This section will explain cryptocurrency manipulation and examine the strategies scammers employ to exploit and endanger investors' investments. After that, we provide an overview of surveys conducted across the scope of price manipulation within the cryptocurrency market and P&D detection within the stock market. Then, we mention valuable studies on P&D prediction in various financial markets. Our analysis will focus on P&D schemes in the context of cryptocurrencies.

2.1. Background

To gain a better understanding of market manipulation, specifically P&D and its effect on cryptocurrencies, this section provides some definitions of the common terms used in this area.

2.1.1. Cryptocurrency Manipulation

Cryptocurrency market manipulation has different aspects. Eigelshoven et al. [3] surveyed papers that have studied cryptocurrency manipulation. They listed seven dif-

ferent methods of cryptocurrency market manipulation, such as P&D [7], wash trading, frontrunning, order book, insider trading, Distributed Denial of Service (DDoS), and stablecoin. Wash trading involves artificially inflating trade volumes by rapid buying and selling, creating the illusion of high market activity to attract investors using multiple accounts and distorting the true value of the asset. It was especially common in less well-known cryptocurrencies and smaller exchanges [8–10]. However, nowadays, with the emergence of non-fungible token (NFT), there are new approaches to wash trading with NFTs [11]. Frontrunning refers to an advantage gained by accessing market information ahead of others; order book refers to creating false price signals by placing and immediately canceling numerous orders; insider trading is the abusive use of privileged information for trading advantages; DDoS refers to the use of distributed denial-of-service attacks to manipulate cryptocurrency markets for profit; stablecoin manipulation refers to suspicious trading practices, potentially affecting cryptocurrency markets. It is worthwhile to add two more types of manipulations: Spoofing is a technique involving placing fake buy or sell orders to create false signals of market activity and liquidity, influencing traders' decisions [12]. Stop hunting is a manipulation tactic involving deliberate actions to trigger the execution of stop-loss orders by driving the price below a critical threshold, creating selling pressure, and providing an opportunity for attackers to buy at lower prices on their analyzed manipulation.

2.1.2. P&D Scheme

In recent years, as displayed in Figure 1, cryptocurrency investors have become increasingly concerned about P&D. There have been too many P&D events in thousands of coins in cryptocurrency, and too many investors have been deceived by manipulators in social media and lost their money by participating in P&D events. Even though those who buy the coin before the pump or even at the beginning of the pump and sell it while it is still pumping have made a profit, many others lose their money because they either entered the event late or held the coin until the price fell below the purchase price.

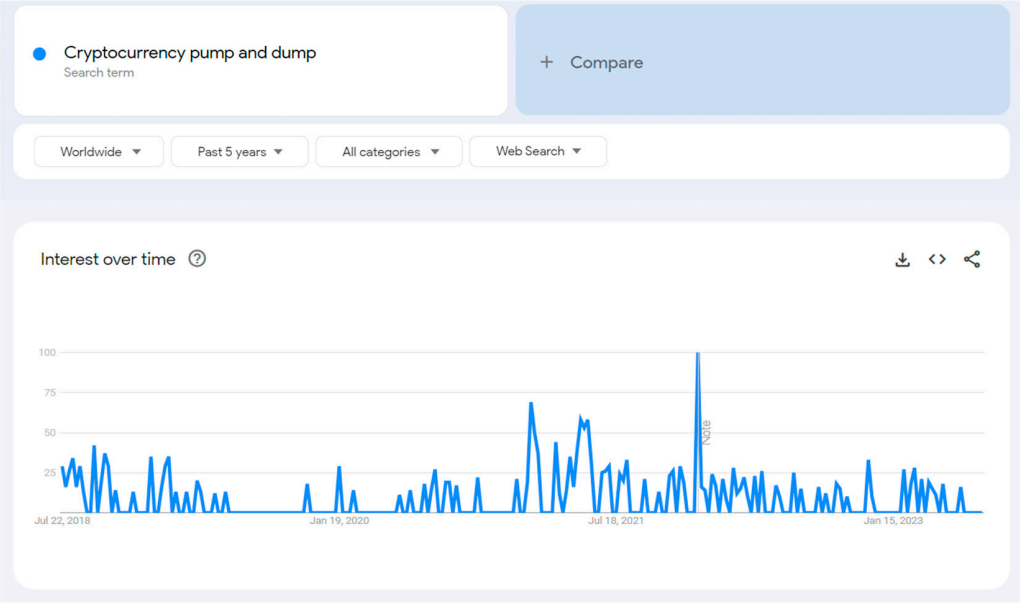

Figure 1. Google trend of cryptocurrency pump and dump.

2.1.3. P&D Phases

Kamps and Kleinberg [13] describe P&D as demonstrated in Figure 2.

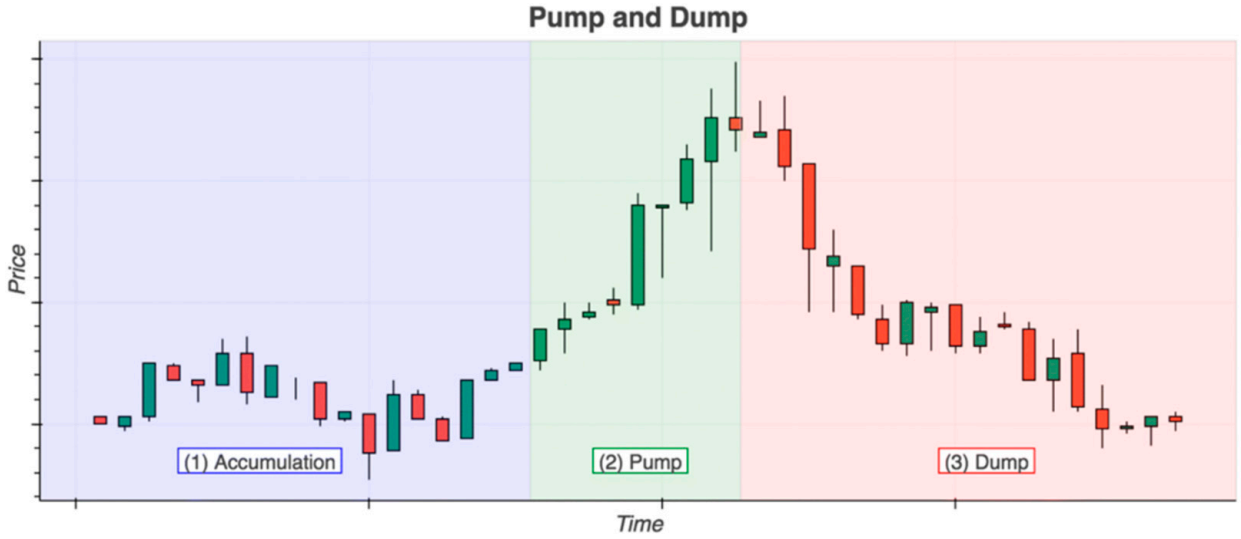

Figure 2. The three phases of P&D [13]. Each color indicates one phase that will be described in the following section.

The P&D event consists of three main phases: accumulation, pump, and finally, dump. In the accumulation phase, the offenders who are looking to create a P&D event start buying the targeted coin in a way that does not lead to an increase in coin price significantly because it may take days to months based on the volume of the coin. Therefore, in this phase, a significant percentage of available coins will be stored in the wallet of those offenders. Then, in the pump phase, the offending use of social media, such as Twitter, Telegram, and Reddit, affects cryptocurrency investors the most. Spreading rumors trying to convince people that the target coin will pump due to multiple factors, for instance, if the coin is not listed in well-known crypto exchanges, they will spread the rumor that it will be listed in Binance [14]. Therefore, some investors will be deceived by this news and start buying that coin, which will inflate the price of the coin. The last phase is a dump in which offenders will start to sell their coins after a profit of several hundred percent, and at one point, they will sell as much as they can, thus creating the peak point. Then, the price will drop below the price before the start of the event. Most of those who bought in the pump phase will be left with useless bags.

There are too many examples of P&D, such as the one mentioned by [15], which happened on BVB coin, or the DATA/BTC in September 2019, mentioned in [16]. We have added one of the most recent P&Ds faced while conducting this survey. This P&D event was organized by a group in Telegram messenger named "Binance Crypto Pumps Signals" [17], with more than 400,000 members. In this group, they organized P&D events in three exchanges of Poloniex [18], Kucoin [19], and Binance. One of the latest P&Ds of this group was on BAR/USDT. They started sending messages in the Telegram group; the first message usually contains data about the date, time, exchange, pairing, and target. After this announcement, they sent some messages about their previous pumps and profits from which members gained to convince members of the channel to participate in this event. They also sent messages clarifying how to buy and sell coins in that specific exchange. Finally, they started to send countdown messages, and, at the announced time, sent a picture that contained the currency name in it. The screenshots are shown in Figure 3. Their last pump occurred on Poloniex Exchange, and the coin was BAR with a pairing of USDT. The chart is shown in Figure 4.

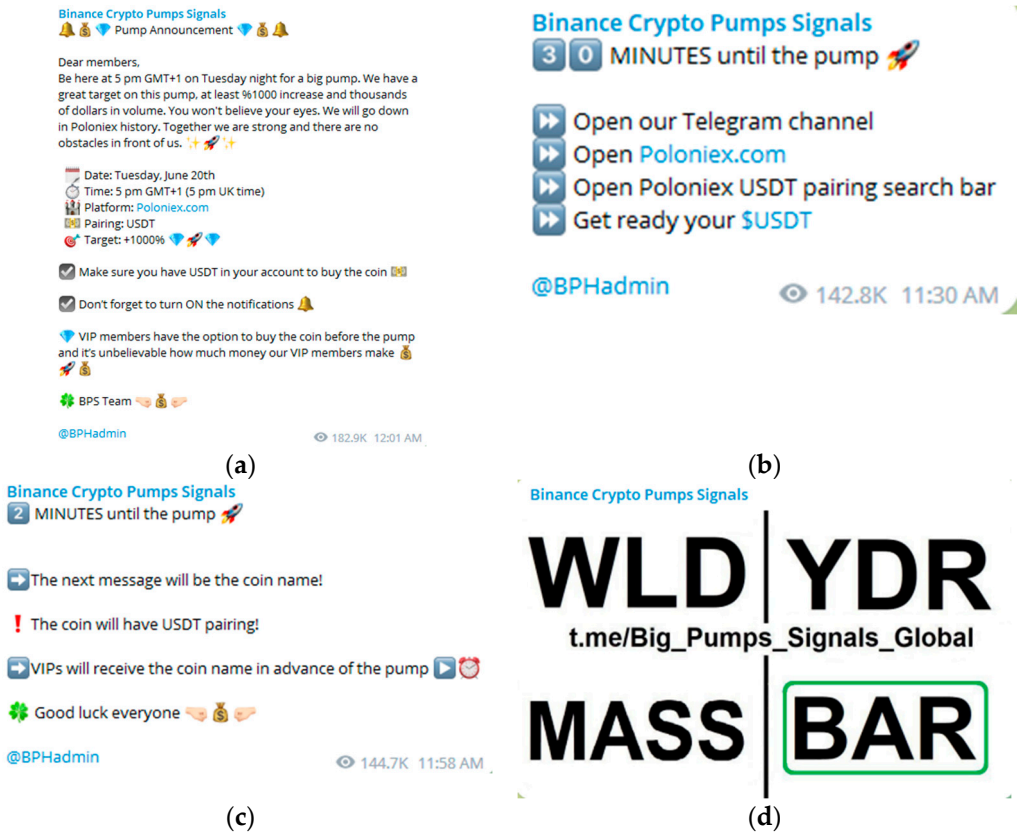

Figure 3. Screenshots of Telegram P&D Group. (**a**) Preparation Announcement; (**b**) Countdown; (**c**) Countdown; (**d**) Coin Announcement (the target coin is shown by green outline).

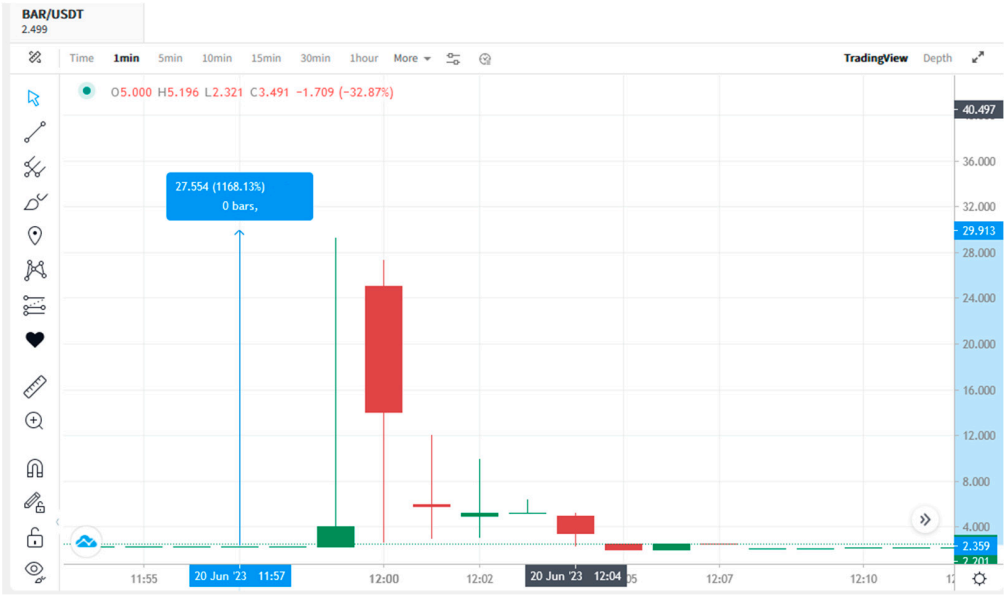

Figure 4. BAR/USDT P&D event organized in a telegram on 20 June 2023. The increase percentage of green candle which indicates pump phase is shown by blue line and the following red candle is dump phase of event. A one-minute time frame was chosen to visualize the candles.

2.2. Related Works

In this section, the related surveys and systematic literature reviews regarding market manipulation detection in stock markets and cryptocurrencies are mentioned. Khodabandehlou and Golpayegani [20] conducted a thorough evaluation of the literature on market

manipulation detection from 2010 to 2020. The authors examined 52 key papers and evaluated definitions, taxonomies, objectives, techniques, strengths and weaknesses, proposed solutions, data mining approaches, and types of manipulations and anomalies. They examined several types of market manipulation including P&D. The study provided several significant outcomes, including realistic solutions for manipulation detection, mapping manipulations with anomalies, identifying obstructions and open questions in the field, and identifying taxonomy criteria for trade-based manipulation. They contributed greatly to showing how electronic financial markets are manipulated and how to detect them. However, they only focused on the manipulations in the stock market, while the crypto market as an electronic financial market was neglected in their study. Furthermore, they indicated an imbalance in focus across various manipulations, with certain manipulations receiving significantly more attention in the studies compared to others. Additionally, they stated that more substantial and in-depth research is needed to develop practical and accurate methods for detecting manipulation. In the survey, it was mentioned that the study's shortcomings include a limited time span (2010–2020), which imply the need for future extended assessments.

Manipulation cases are increasingly complex, making it impossible to observe them manually. To detect them, intelligent systems have been developed. Zulkifley et al. [21] conducted a review that focused on methods and algorithm approaches in conventional machine learning (ML) and deep learning (DL). In their report, they stressed the importance of well-managed financial markets and the role that regulators play in preventing illegal trading. They also emphasized on the effectiveness of using artificial intelligence in detecting manipulations, particularly the preference for tick data over daily trading data when detecting manipulations. It turned out that support vector machine (SVM) classifiers were the most popular, and the detection rates were often higher with simpler classes. Furthermore, the paper outlined challenges facing researchers, such as regulatory differences, a lack of publicly available datasets and manipulation case listings, and the need for algorithm performance improvements. To enhance detection capabilities while addressing regulatory and data challenges, the authors suggested implementing DL analysis networks as well as incorporating advanced attention mechanisms and stock data from multiple countries. Nevertheless, this study also focused on detecting manipulations in stock markets and neglected the crypto market.

Regarding the cryptocurrency markets, Eigelshoven et al. [3] conducted a systematic literature review to provide a comprehensive overview of cryptocurrency market manipulation methods. Seven manipulation methods in cryptocurrency markets are listed in their study, while six of them were already known from traditional markets. However, they stated that there is a gap between academic research and market activities; there may still be current market manipulation techniques that have not yet received academic attention. Furthermore, they identified six factors that lead to successful market manipulation schemes. In the absence of regulation and standard know-your-customer (KYC) procedures, exchanges play an important role in market manipulation. In the study, the authors suggest that using the identified market vulnerabilities to develop regulation approaches can be beneficial. Nevertheless, the study has some limitations, including a possible omission of important studies and subjective selection of results.

Lastly, Table 1 summarizes the main scopes of these three studies compared to this survey. Accordingly, Khodabandehlou and Golpayegani [20] and Zulkifley et al. [21] focused on detecting manipulations only in the stock markets. However, our survey aims to focus on the studies that detected a unique type of market manipulation called P&D in the cryptocurrency markets. Moreover, although Eigelshoven et al. [3] provided an overview of market manipulations in cryptocurrency, including P&D, they did not focus on the methods to detect these manipulations. On the other hand, our study focused on the methods used to detect P&D.

Table 1. Summary of recent surveys and comparison with this review.

Paper Title	Market Manipulation	P&D	ML for Detection	Cryptocurrency	Summary
Cryptocurrency Market Manipulation: A Systematic Literature Review [3]	yes	yes	no	yes	A comprehensive survey on cryptocurrency manipulation papers that provides a complete definition of different manipulations in cryptocurrency and identifies market vulnerabilities.
Market manipulation detection: A systematic literature review [20]	yes	yes	yes	no	A survey of the literature on market manipulation detection from 2010 to 2020. It identifies different manipulations and focuses on trade-based manipulation.
A Survey on Stock Market Manipulation Detectors Using Artificial Intelligence [21]	yes	yes	yes	no	A survey that aims to discuss state-of-the-art automated methods for detecting manipulations. It also defines a manipulation taxonomy.
This Survey	yes	yes	yes	yes	A comprehensive survey that examines the recent progress in using ML to detect and predict P&D in the cryptocurrency market.

Besides studies addressing P&D detection in cryptocurrency markets, studies have been conducted across various financial markets addressing market manipulation detection in a broader context. Several approaches have been explored, including supervised and unsupervised DL techniques [22,23], transformer models [24], immune-inspired dendritic cell algorithms [25,26], Markov models [27], and other ML approaches [28–33]. Models trained with supervised DL have been used to detect and classify manipulative trading behavior, whereas models trained with unsupervised DL focus on clustering and anomaly detection. Transformer models leverage self-attention mechanisms to capture dependencies between trading events, while immune-inspired dendritic cell algorithms mimic the immune system to detect anomalous activities. P&D detection techniques in cryptocurrency markets can become more comprehensive and effective by incorporating insights from these approaches or using transfer learning to finetune the previous models for detecting specific market manipulations like P&D.

3. Methodology

Information on our review technique is provided in this section, including our search strategies and sources, data collection, and extraction strategy.

3.1. Search Strategies

In this survey, our focus was on surveying approaches to detect P&D in the cryptocurrency market. To ensure a comprehensive examination of the literature, we used different terms that best describe the P&D and price manipulation in cryptocurrency. These strategies entailed the utilization of pertinent keywords associated with fraud detection and cryptocurrency, such as "Pump and Dump" and "cryptocurrency manipulation." The search was carried out using the keywords below:

- ("Pump and Dump" AND "Twitter"),
- ("Pump and Dump " AND "Machine learning''),
- (Real AND time AND cryptocurrency AND market AND manipulations),
- ("Pump and Dump" AND "deep learning"),
- ("Pump and Dump" AND "cryptocurrency"),
- (Cryptocurrency AND manipulation AND prediction).

3.2. Search Sources

We conducted our searches using academic databases of high repute, such as Scopus and Google Scholar. This selection was based on the extensive coverage of scholarly publications across diverse disciplines, such as finance, economics, and computer science. Multiple search sources were incorporated to minimize the possibility of overlooking relevant papers.

3.3. Inclusion and Exclusion Criteria

The quality and relevance of the survey were ensured by selecting the papers that should be included and excluded. For the first step, articles discussing P&D schemes in cryptocurrencies were considered. We placed a special emphasis on papers utilizing ML models to detect P&D schemes. Furthermore, only English language papers addressing market manipulation in cryptocurrencies were considered for the selection. Table 2 provides a summary of the inclusion and exclusion criteria that we have employed.

Table 2. Inclusion exclusion criteria.

Inclusion Criteria	Exclusion Criteria
English	Languages other than English
P&D	Other forms of manipulation
Crypto currencies	Other stock markets
Utilizing ML for P&D detection	Absence of model for P&D detection

3.4. Data Collection Procedure

We conducted the selection process in May 2023. The process of selecting papers is shown in Figure 5.

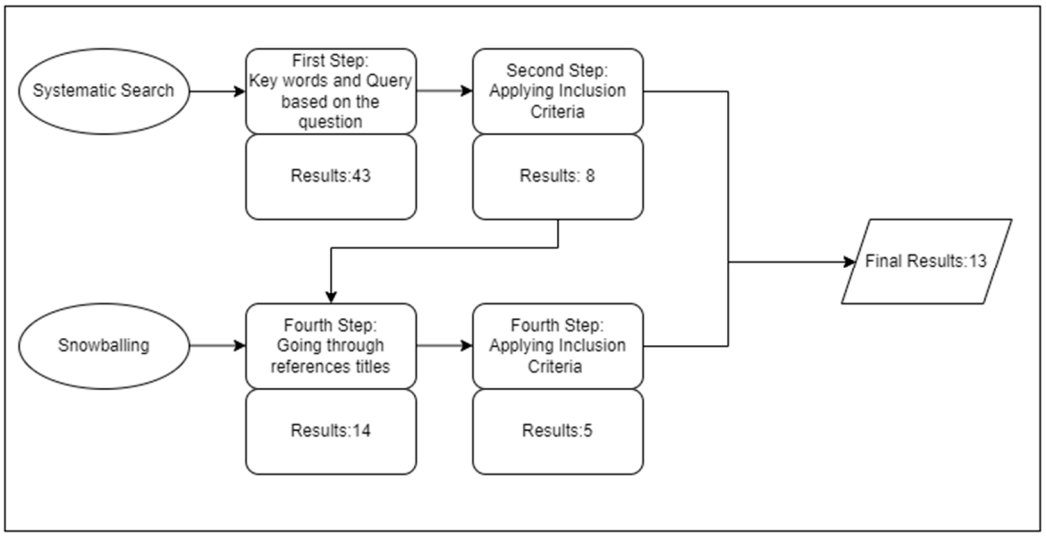

Figure 5. Collection procedure.

We employed a comprehensive two-step data collection procedure, including systematic searching and snowballing which involves expanding the collected papers by searching their references to find more related papers not originally retrieved [34]. As a result, we ensured that a wide range of relevant papers were reviewed. To begin with, a systematic search was conducted to retrieve a core set of papers. Utilizing a technique called snowballing, additional related papers were found by examining the references of these papers.

The first phase included a systematic search on Scopus and Google Scholar using predefined search queries. Through the search process, 43 papers were initially retrieved. We then screened these papers based on predefined inclusion criteria. As a result of applying our inclusion criteria, eight papers met the requirements for our survey, while 35 papers were excluded. Our first phase selection consisted of eight papers.

For the second step, we reviewed all the selected papers' reference titles. As a result of this process, 14 additional papers were identified as relevant to our study. After that, we applied our inclusion criteria to each of these newly identified papers, and those that fell outside our inclusion criteria were excluded. Consequently, 5 additional papers were included, bringing the total number of selected papers for our survey to 13.

3.5. Data Extraction Strategy

A structured data extraction strategy was implemented to extract key information from each selected paper. This encompassed details such as author information, publication year, research methodology, utilized datasets, ML techniques employed, performance metrics, and noteworthy findings.

Through the collected data, this survey endeavors to provide a comprehensive overview of the research conducted in the realm of P&D detection in cryptocurrency to address the research question, which was the following: How to analyze social media and market data to detect P&D using ML?

4. Findings

Based on the survey methodology and data collection, 13 papers on P&D detection using ML in cryptocurrency were found. In this section, in the form of a complete literature review, we go through selected papers and mention models that have been deployed to detect P&D and their results. Moreover, a brief description of all the studies that were mentioned in this survey is provided in Table 3. Each study used a different data collection and labeling policy, which affected their training approaches. Considering the specific training methodologies used in each study, we have categorized these studies into three groups: supervised approach, semi-supervised approach, and unsupervised approach.

4.1. Supervised Approach

Several studies provided manually labeled datasets with confirmed pump events to train ML and DL models that can detect the start of P&D. First, we explored the studies that employed classic ML to detect P&D. For instance, Xu and Livshits [15] analyzed P&D activities organized in Telegram channels between June 2018 and February 2019 and developed several models, such as random forest (RF) and a generalized linear model (GLM), to predict the likelihood of a pump for all listed coins on a crypto exchange. Their training data consisted of 60 true samples and 20,000 false samples. Since the true/false ratio was 0.3%, they decided to decrease the number of false samples and increase the number of trees in RF; they trained three models (RF1, RF2, RF3) with 20,000, 5000, and 1000 false samples and 5000, 10,000, and 20,000 number of trees, respectively. Since the distribution of positive samples is uneven, conventional binomial GLM would be insufficient; therefore, this paper applied least absolute shrinkage and selection operator (LASSO) regularization to the GLM models. The results showed that RF models performed better in terms of precision and F1-score, where RF3 outperformed all RF models with a low threshold (0.2). Furthermore, Shao [35] conducted a study to analyze the effectiveness of modern data science techniques in detecting fraudulent activities in the Dogecoin market. He highlighted the need for objective detection of fraudulent activities and the consideration of the influence of factors, like Elon Musk's tweets, on cryptocurrency prices. He utilized models such as RF, LR, SVM, k nearest neighbors (KNN), and decision trees (DT) combined with bagging and stacking techniques. The RF algorithm, trained with a 5-fold cross validation (CV) technique, achieved the best out-of-sample testing accuracy with 100%. Moreover, the good performance of RF in F1, precision, and recall scores mitigated concerns about overfitting.

La Morgia et al. [36] suggested several features to detect the start of the P&D procedure. The data were split into chunks of "S" seconds and a moving window of W hours. It is necessary to choose small values for chunk size (S) to detect the P&D procedure in its early stages. They used nine features for classification: StdRushOrders, which involves the moving standard deviation of the volume of rush orders in each chunk; AvgRushOrders, which involves the moving average of the volume of rush orders in each chunk; StdTrades, which involves the moving standard deviation of the number of buys and sell trades; StdVolumes, which involves the moving standard deviation of the volume of trades in each chunk; AvgVolumes, which involves the moving average of the volume of trades in each chunk; StdPrice, which involves the moving standard deviation of the closing price;

AvgPrice, which involves the moving average of the closing price; AvgPrice-Max, which involves the moving average of the maximal price in each chunk; AvgPriceMin, which involves the moving average of the minimal price in each chunk. They used RF and logistic regression (LR) for the classification task. Each RF model consisted of 200 trees with a max depth of four. In all the models that were trained for different chunk sizes, RF showed slightly better performance than LR. The results showed that a chunk size of 25 s and a window size of 7 h achieved the best F1-score, and a chunk size of 5 s and a window size of 50 min had the best speed. Furthermore, in another study [37], the authors used the features from their previous study and several new features, such as HourSin, HourCos, MinuteCos, and MinuteSin, where the hour and minute of the first transaction in each chunk is encoded with sine and cosine functions. This study used an RF model with 200 trees and a max depth of five for each tree. The RF model showed outstanding precision results, but after decreasing the chunk size from 25 to 5, recall dropped significantly. To solve this problem, they suggested utilizing the AdaBoost model; this model showed stable precision and recall and an improved F1-score. Also, the results showed that increasing the chunk size causes an improvement of the F1-score.

All the mentioned studies have used classic ML algorithms to detect P&D. However; some studies employed modern approaches, such as DL models, to predict P&D schemes. For instance, Chadalapaka et al. [38], used the same input features as [37] in addition to two other features, namely pump index and symbol. They utilized CLSTM and an anomaly transformer as novel approaches to detect P&D schemes in cryptocurrencies. The CLSTM model is made of convolutional layers which find the spatial features; max-pooling layers; LSTM layers which find the temporal features; and finally, fully connected layers for final prediction. The anomaly transformer introduces two novelties: an anomaly attention module and a minimax optimization strategy. The anomaly attention module computes series association and prior association. Minimax optimization consists of two phases: during the minimize phase, the series association moves toward the prior association, and during the maximize phase, the series association moves toward the original input. The results demonstrated that these DL models outperform classic ML models and statistical methods for detecting such fraudulent activities, where the anomaly transformer with a chunk size of 15 s and F1-score of 93.6 showed the best performance. In another study, Nghiem et al. [16] used market data consisting of OHLCV data and social data, which were collected from official Reddit accounts, Facebook pages, and Twitter accounts to train four types of models that detect P&D. Training, validation, and test sets contained 197, 55, and 54 pump events, respectively. They trained the models in different training situations in terms of the type of training data (social, financial, or both) and lookback time window. These models consisted of baseline (LR), CNN, Bidirectional LSTM, and CLSTM. The CNN model is a convolutional neural network that is widely used in DL and is powerful in feature detection. Bidirectional LSTM finds sequential dependencies in both directions in an input sequence. CLSTM demonstrated the most promising performance and consistency, while BLSTM models exhibited inconsistent performance across configurations.

Besides DL models, there are other modern approaches, such as XGBoost, that can be utilized in the literature. For instance, Victor and Hagemann [39] utilized BTC trading pairs data from Binance and chat histories from Telegram P&D groups to label the data. They examined the characteristics of pump announcements, the timing of pumps, simultaneous announcements across channels, and the price action following a pump. Three types of pump events were identified: sustained pumps, short-term P&D events, and failed pumps. They developed an XGBoost model to detect P&D events automatically, and they achieved a high AUC of 0.995, demonstrating good sensitivity and specificity on the test set. They created a dataset with positive and negative ground truth samples and engineered features, such as volume bars, price change rates, and time differences. Applying the model to 172 coins, it identified 612 pump-like events, primarily in low-market capitalization cryptocurrencies, confirming its effectiveness in detecting suspicious activity. The unique

aspects of this study are high-resolution data analysis, focusing on Binance, to explore profitability, medium-term impact, and the utilization of ground truth for detection.

4.2. Semi-Supervised Approach

Since there is a large number of telegram messages and tweets, manually labeling these messages to find whether they are pump related or not is time-consuming and expensive. To address this issue, some studies deployed a semi-supervised approach to label the messages. They manually labeled some of the messages, and then they trained a model to predict the labels; they used this trained model to label all the messages. Mirtaheri et al. [40] combined information from Twitter and Telegram to automatically detect ongoing pump operations and predict their success. They labeled 1557 messages and trained an SVM to label their data. They used three kinds of data for detecting the pump attempts: Telegram data which contain 195,576 Telegram messages; Twitter data which includes 3,760,831 tweets; and market data which are collected from the CoinMarketCap website [41]. They trained two binary classification models for detecting pump attempts from Telegram messages and to determine whether the target price would be achieved or not. The results of their ML models showed reasonable accuracy in detecting unfolding attacks and predicting the achievement of price targets. They also found that these scams are often accompanied by an increase in bot activities on Twitter.

On the other hand, Hu et al. [42] labeled 5050 messages and used RF and LR models to label the data. They trained DL models to classify the input data and determine whether the input was pump-related or not. They focused on predicting the pump probability of coins listed in a specific exchange before the scheduled pump times. The authors analyzed 709 P&D events organized in Telegram from January 2019 to January 2022 and identified interesting patterns, such as intra-channel homogeneity and inter-channel heterogeneity among pumped coins. Inspired by these observations, they developed a novel sequence-based neural network called SNN which mainly consists of three parts: an embedding layer that is used to handle categorical values in the input sequence and reduce sparsity and dimensionality by converting the sequence into a dense vector representation; positional attention that is used to find the sequential patterns in the input sequence and capture skip-correlations; and a multi-layer perceptron containing several fully connected layers to predict the final pump probability. Extensive experiments validated the effectiveness and generalizability of their proposed methods. Overall, their data science pipeline and SNN model showed promising results, indicating their suitability for real-world scenarios.

4.3. Unsupervised Approach

Some of the studies focused on finding abnormalities in the input data and classifying them as possible P&D schemes while they are not associated with labeled data for training. For instance, Chen et al. [43] proposed an algorithm to detect user groups that are potentially involved in P&D schemes. To validate their algorithm, they analyzed the leaked transaction history of the Mt. Gox Bitcoin exchange, and, using an improved Apriori algorithm, they identified several user groups engaging in synchronized buying or selling. This data contained 14 million records and 120,000 users with information about user ID, date, type, and the amount of each transaction. The authors discovered numerous abnormal trading behaviors and pricing patterns within these detected groups. For example, some users buy much more Bitcoin than they sell, and in reverse. There are also a considerable number of users that never buy or sell Bitcoin in the system. The authors suggested that some of these abnormalities could be linked to accounts controlled by the exchange itself for some special reason, for example, balancing the trading and providing liquidity.

Some studies used anomaly detection to find abnormal patterns in the input data and mark them as possible P&D schemes. Kamps and Kleinberg [13] were the first ones to propose criteria to define a cryptocurrency P&D and utilized anomaly detection techniques to identify anomalous trading activities that could indicate potential P&D schemes. They used OHLCV data from symbol pairs obtained from five different exchanges through the

CCXT library [44]. This study detected three kinds of anomalies to analyze the data: price anomaly, volume anomaly, and pump anomaly, which is defined as a co-occurrence of previous anomalies. They detected anomalies with three different sets of parameters. The results indicated that there are signals in the trading data that can help detect these fraudulent activities. This study suggested that more regulated exchanges are less vulnerable to P&D schemes. Furthermore, Mansourifar et al. [45] proposed a hybrid anomaly detection based on distance and density metrics to address the challenge of anomaly detection in time series data. An automatic threshold-setting method was developed by the authors for distance-based anomaly detection, as well as the introduction of a new metric called the density score for density-based anomaly detection. In conclusion, the authors discussed the challenges of contextual anomaly detection and outlined their approach to transforming the problem into a point anomaly detection task. According to their experiments, the hybrid approach outperformed both density-based and distance-based approaches at detecting P&D in top-ranked exchange pairs.

Previous studies with unsupervised approaches used classic ML algorithms to detect possible P&D cases. DL can also be used to detect these cases. Bello et al. [46] trained an LSTM-based auto-encoder on Bitcoin valuations using a low latency detection (LLD) framework based on DL. The auto-encoder was then used to predict valuations on altcoins. Since this study uses an unsupervised approach, they did not need ground truth labels for training the model, but for testing the model, they collected 55 confirmed pump events from Telegram groups to make this data relevant. They also retrieved OHLCV data and the number of trades as input features to the model. This model was trained on the data obtained one day before each of the 55 pumps. They also trained this model on the data from two and three days before the pumps, and the results showed that the model that was trained with data from one day before the pumps had the best performance compared to the other two models. Finally, they performed anomaly detection on the predicted and actual prices to find the anomalous points. The study highlighted the advantages of LLD, such as practical deployment, unsupervised training, and fast inference, while eliminating the need for future data.

Table 3. Brief description of the studies mentioned in this survey.

Paper	Number of Coins	Exchanges	ML Models	Criteria	Best ML
Kamps and Kleinberg [13]	More than 50 pairs	Binance, Bittrex, Kraken, Kucoin, Lbank	Anomaly detection	Accuracy	Anomaly detection
Victor and Hagemann [39]	172	Binance	XGBoost	Sensitivity, specificity	XGBoost
Xu and Livshits [15]	296	Binance, Bittrex, Cryptopia, Yobit	RF, GLM	F1, AUC, precision	RF1
Chen et al. [43]	1	Mt. Gox	Apriori	N/M	Improved a priori algorithm
Morgia et al. [36]	194	Binance	RF and LR	Precision, recall, F1	RF (10 folds) with chunk size 25 S
Mansourifar et al. [45]	10	Lbank, Kucoin, Bittrex, Binance	Anomaly detection	Accuracy	Anomaly detection
Nghiem et al. [16]	355	Binance, Bittrex, Cryptopia, Yobit	LR as baseline model, CNN, BLSTM, and CLSTM	MAPE, precision, recall, F1	CNN Fin 6
Mirtaheri et al. [40]	543	N/M	SVM	Accuracy, precision, recall, F1	SVM
Shao [35]	1	Binance	DT + CV, RF + CV, LR + CV, SVM + CV, and an ensemble of LR, RF, and SVM	Accuracy, F1, precision, recall	RF(5-fold)
Chadalapaka et al. [38]	194	Binance	CLSTM, Anomaly Transformer	Precision, recall, F1	Anomaly transformer

Table 3. *Cont.*

Paper	Number of Coins	Exchanges	ML Models	Criteria	Best ML
Hu et al. [42]	1	Binance	LR, RF, DNN, LSTM, BLSTM, GRU, BGRU, TCN, SNN	AUC, precision, recall, F1, heat ratio	SNN
Morgia et al. [37]	378	Binance	RF, AdaBoost	F1, recall, precision	AdaBoost
Bello et al. [46]	Coins with a pair of BTC and a few USDT and ETH as well	Binance	LSTM-based auto-encoder	Precision, recall, F1	LSTM-based auto-encoder

5. Discussion and Research Directions

This section delves into the critical aspects of our findings and analysis. Section 5.1 discusses the various methodologies employed to detect P&D, and Section 5.2 outlines the research directions and discusses potential approaches and solutions.

5.1. Discussion

The purpose of this section is to present an overview of the 13 studies that investigated the use of ML techniques in order to predict P&D manipulations in cryptocurrencies. Table 4 presents the strengths and weaknesses of these studies. Classical ML models were initially adopted by researchers because of their ease of implementation and the ability to achieve satisfactory performance on smaller datasets without overfitting. Subsequently, various studies explored the application of unsupervised ML approaches [13,43,45], while others employed supervised ML techniques, such as RF or LR [15,35,36,40]. Additionally, more recent investigations have utilized contemporary machine learning models, like XGBoost and AdaBoost [37,39]. As research in this domain progressed, it facilitated the development of larger datasets and the training of DL models, which have become prevalent in addressing modern problem-solving challenges [16,38,42,46]

Table 4. Strengths and limitations of each study.

Paper	Strengths	Limitations
Kamps and Kleinberg [13]	- Primitive academic research on cryptocurrency P&D schemes - Unsupervised detection of P&D schemes	- Limited dataset (20 days with hourly granularity) - Lack of a confirmed P&D database
Victor and Hagemann [39]	- Labeling pump events based on Telegram messages	- Lack of a sufficient number of ML models for proper comparison of results - Limited scope: Analysis just on one exchange
Xu and Livshits [15]	- Extracting effective features for the input data	- Lack of a confirmed P&D database - Imbalanced data - Low performance of the models
Chen et al. [43]	- Detecting groups of traders that buy/sell together to detect possible P&D schemes	- Limited scope: analysis restricted to a single cryptocurrency (BTC) - The model cannot be utilized to detect real-time P&D in the open market
Morgia et al. [36]	- Defining rush orders as a feature for input data	- Limited scope: analysis just on one exchange - Insufficient split size for test-set

Table 4. *Cont.*

Paper	Strengths	Limitations
Mansourifar et al. [45]	- Proposing a novel approach using a combination of distance and density metrics	- Lack of a confirmed P&D database
Nghiem et al. [16]	- Providing the data for both social data and market data - Statistical analysis, pairwise *t*-tests	- Low performance of the proposed models
Mirtaheri et al. [40]	- A large number of data from telegram, Twitter, and market	- Increased uncertainty of the pump detection model due to weakly supervised labeling of the data - Lack of a sufficient number of ML models for proper comparison of results
Shao [35]	- Implementing various models and employing various techniques to improve the training to achieve better results	- Lack of a confirmed P&D database—The analysis is exclusive to Elon Musk tweets and Dogecoin market data from a single exchange
Chadalapaka et al. [38]	- Introducing novel DL approaches for detecting sequential anomalies with good performance	- Limited scope: analysis just on one exchange
Hu et al. [42]	- Introducing an attention-based approach that captures skip correlations to detect pump attempts	- Increased uncertainty of the pump detection model due to weakly supervised labeling of the telegram data
Morgia et al. [37]	- Defining hour and minute features encoded with cosine and sine for input data	- Limited scope: analysis just on one exchange; insufficient split size for test-set
Bello et al. [46]	- Introducing an unsupervised LSTM-based autoencoder to detect pump events	- Lack of a sufficient number of ML models for proper comparison of results - Limited scope: analysis just on one exchange

Many of these studies lack a reliable P&D database. It is difficult to identify P&D instances without a dataset containing confirmed P&D events. There were limitations due to limited data availability in several studies [13,15,16,35–40,42,43,45,46]. A large amount of data are required for effective training of ML models. It is possible that insufficient and imbalanced data result in training models that lead to errors when predicting new data, which could potentially have a negative impact on the authors' conclusions and arguments. In addition, it is possible for models performing well with small datasets not to replicate the same results when applied to larger and more representative samples. As a result, the model's performance should always be evaluated using a substantial, adequate, and valid dataset. There is also a notable limitation in some studies for utilizing a limited number of models to address the issue [15,36,37,39,46]. Using only one model to predict P&D activities makes it difficult to rely on the results since there is no benchmark for comparing the model's performance. The use of multiple models for tackling a problem is not only beneficial in comparing their performance but also enhances confidence in their effectiveness.

In addition, the use of a limited number of performance analysis criteria represents another limitation [43,45]. The performance of a model cannot be adequately assessed by relying on a single metric. Using multiple criteria to evaluate the model's performance allows for more robust conclusions to be drawn. A further limitation is studying a limited

number of cryptocurrencies or using only data from one cryptocurrency exchange [35,42,43], which might result in training models that detect patterns specific to a single cryptocurrency.

Our survey reveals that various articles have utilized different approaches to detect pump and dump events, incorporating factors such as candle prices, trading volume, and social media data. To defend against pump and dump, one important step would be to educate cryptocurrency traders about the taxonomy of pump and dump schemes. By increasing their awareness, traders can be more cautious and less likely to fall into these traps. Additionally, the approaches highlighted in our survey can be employed to predict the likelihood of a coin being subjected to pump and dump, providing traders with valuable insights to make more informed decisions and mitigate potential risks.

5.2. Research Directions

The following research directions can be explored to advance P&D detection in cryptocurrency:

- Detecting the end of the pump phase: It would be beneficial to develop methods for detecting the end of the pump phase with a high degree of accuracy. The unavailability of this information on social media platforms makes accurate detection difficult. To address this future work, researchers can explore advanced ML algorithms that can analyze historical price patterns and social media sentiments to detect abrupt changes in market behavior, indicating the end of the pump phase;
- Fake news and coordinated advice impact: One possible future direction is to consider the factor of fake news and coordinated investment advice. Since many individual investors rely on investment websites, news channels, and investment advice channels, if enough of them buy a specific coin because of fake news or coordinated advice, it will cause the price to change quickly. To tackle this research direction, sentiment analysis can be employed to identify false information and advice. Additionally, developing a reliable system that verifies the authenticity and credibility of news sources could help mitigate the influence of false information on the market;
- Feature combinations for enhanced model training: Various combinations of input features in the data can be used to train the models, for example, by using indicators as input features for market data, such as RSI, which considers whether an asset is overbought or oversold; MACD, which tries to forecast market trends by comparing short and long-term tendencies; SMA and EMA, which represent simple and exponential moving averages of the market data, respectively. These indicators are used to improve the efficiency of the prediction process and develop the quality of the data, which could be beneficial to the performance of the trained models. Incorporating and examining the outcomes of different models using such combinations of input data could be the subject of future studies aiming to establish general guidelines for future research in this field. Researchers can address this future work by conducting an extensive feature engineering analysis to identify the most relevant indicators and combine them to enhance model training;
- Transfer learning: There are studies [20–22] that train ML models for market manipulation detection; future studies can use transfer learning to fine-tune these pre-trained models for detecting specific market manipulations like P&D.

6. Conclusions

The survey provided a comprehensive overview of the academic literature on P&D detection in cryptocurrency. Conducting this research was necessary due to the fact that P&D manipulation has become a serious challenge for cryptocurrency market participants, followed by the increasing popularity of cryptocurrency, the number of coins, types of blockchains, and the complexity of manipulation methods. Therefore, several studies have addressed the problem of P&D detection using different approaches. This study has provided guidance for researchers to utilize and optimize these methods to enhance P&D detection. Most of the studies that we mentioned deployed classic ML models to analyze

the market and social data. Among these models, RF showed a better performance. Several of the studies deployed DL models that performed even better than classic ML models. However, as there is a large number of trainable variables, these models need more training data to be reliable.

Author Contributions: Writing—original draft preparation, M.J.R.; supervision and writing—review and editing, Q.H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liang, J.; Li, L.; Chen, W.; Zeng, D. Towards an understanding of cryptocurrency: A comparative analysis of cryptocurrency, foreign exchange, and stock. In Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics (ISI), Shenzhen, China, 1–3 July 2019; pp. 137–139.
2. Bunjaku, F.; Gjorgieva-Trajkovska, O.; Miteva-Kacarski, E. Cryptocurrencies–advantages and disadvantages. *J. Econ.* **2017**, *2*, 31–39.
3. Eigelshoven, F.; Ullrich, A.; Parry, D.A. Cryptocurrency market manipulation: A systematic literature review. In Proceedings of the 42nd International Conference on Information Systems, Austin, TX, USA, 12–15 December 2021.
4. Baum, S.C. Cryptocurrency Fraud: A Look into the Frontier of Fraud. Bachelor's Thesis, Georgia Southern University, Statesboro, GA, USA, 2018.
5. Alyami, M.; Alhotaylah, R.; Alshehri, S.; Alghamdi, A. Phishing Attacks on Cryptocurrency Investors in the Arab States of the Gulf. *J. Risk Financ. Manag.* **2023**, *16*, 271. [CrossRef]
6. Bonifazi, G.; Corradini, E.; Ursino, D.; Virgili, L. A social network analysis–based approach to investigate user behaviour during a cryptocurrency speculative bubble. *J. Inf. Sci.* **2021**, *49*, 1060–1085. [CrossRef]
7. Hamrick, J.; Rouhi, F.; Mukherjee, A.; Feder, A.; Gandal, N.; Moore, T.; Vasek, M. An examination of the cryptocurrency pump-and-dump ecosystem. *Inf. Process. Manag.* **2021**, *58*, 102506. [CrossRef]
8. Victor, F.; Weintraud, A.M. Detecting and quantifying wash trading on decentralized cryptocurrency exchanges. In Proceedings of the Web Conference 2021, Ljubljana, Slovenia, 19–23 April 2021; International World Wide Web Conference Committee: Ljubljana, Slovenia, 2021; pp. 23–32.
9. Le Pennec, G.; Fiedler, I.; Ante, L. Wash trading at cryptocurrency exchanges. *Financ. Res. Lett.* **2021**, *43*, 101982. [CrossRef]
10. Cong, L.W.; Li, X.; Tang, K.; Yang, Y. *Crypto Wash Trading*; National Bureau of Economic Research: Washington, DC, USA, 2022.
11. Bonifazi, G.; Cauteruccio, F.; Corradini, E.; Marchetti, M.; Montella, D.; Scarponi, S.; Ursino, D.; Virgili, L. Performing Wash Trading on NFTs: Is the Game Worth the Candle? *Big Data Cogn. Comput.* **2023**, *7*, 38. [CrossRef]
12. Cartea, Á.; Jaimungal, S.; Wang, Y. Spoofing and price manipulation in order-driven markets. *Appl. Math. Financ.* **2020**, *27*, 67–98. [CrossRef]
13. Kamps, J.; Kleinberg, B. To the moon: Defining and detecting cryptocurrency pump-and-dumps. *Crime Sci.* **2018**, *7*, 18. [CrossRef]
14. Binance. Public Rest Api for Binance. Available online: https://github.com/binance-exchange/binance-official-api-docs/ (accessed on 20 June 2023).
15. Xu, J.; Livshits, B. The Anatomy of a Cryptocurrency Pump-and-Dump Scheme. In Proceedings of the USENIX Security Symposium, Santa Clara, CA, USA, 14–16 August 2019; pp. 1609–1625.
16. Nghiem, H.; Muric, G.; Morstatter, F.; Ferrara, E. Detecting cryptocurrency pump-and-dump frauds using market and social signals. *Expert Syst. Appl.* **2021**, *182*, 115284. [CrossRef]
17. Big Pumps Signals Global. Available online: https://t.me/Big_Pumps_Signals_Global (accessed on 26 June 2023).
18. Poloniex. Available online: https://poloniex.com/ (accessed on 23 June 2023).
19. kucoin. Available online: https://www.kucoin.com/ (accessed on 23 June 2023).
20. Khodabandehlou, S.; Golpayegani, S.A.H. Market manipulation detection: A systematic literature review. *Expert Syst. Appl.* **2022**, *210*, 118330. [CrossRef]
21. Zulkifley, M.A.; Munir, A.F.; Sukor, A.; Edil, M.; Mohd Shafiai, M.H. A Survey on Stock Market Manipulation Detectors Using Artificial Intelligence. *Comput. Mater. Contin.* **2023**, *75*, 4395–4418.
22. Chullamonthon, P.; Tangamchit, P. Ensemble of supervised and unsupervised deep neural networks for stock price manipulation detection. *Expert Syst. Appl.* **2023**, *220*, 119698. [CrossRef]
23. Leangarun, T.; Tangamchit, P.; Thajchayapong, S. Stock price manipulation detection using deep unsupervised learning: The case of Thailand. *IEEE Access* **2021**, *9*, 106824–106838. [CrossRef]
24. Chullamonthon, P.; Tangamchit, P. A transformer model for stock price manipulation detection in the stock exchange of Thailand. In Proceedings of the 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Prachuap Khiri Khan, Thailand, 24–27 May 2022; pp. 1–4.

25. Rizvi, B.; Belatreche, A.; Bouridane, A. A dendritic cell immune system inspired approach for stock market manipulation detection. In Proceedings of the 2019 IEEE Congress on Evolutionary Computation (CEC), Wellington, New Zealand, 10–13 June 2019; pp. 3325–3332.
26. Rizvi, B.; Belatreche, A.; Bouridane, A. Immune inspired dendritic cell algorithm for stock price manipulation detection. In *Intelligent Systems and Applications: Proceedings of the 2019 Intelligent Systems Conference (IntelliSys)*; Springer: Cham, Switzerland, 2020; Volume 1, pp. 352–361.
27. Cao, Y.; Li, Y.; Coleman, S.; Belatreche, A.; McGinnity, T.M. Adaptive hidden Markov model with anomaly states for price manipulation detection. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *26*, 318–330. [CrossRef]
28. Uslu, N.C.; Akal, F. A machine learning approach to detection of trade-based manipulations in Borsa Istanbul. *Comput. Econ.* **2022**, *60*, 25–45. [CrossRef]
29. Yagemann, C.; Chung, S.P.; Uzun, E.; Ragam, S.; Saltaformaggio, B.; Lee, W. On the feasibility of automating stock market manipulation. In Proceedings of the Annual Computer Security Applications Conference, Austin, TX, USA, 7–11 December 2020; pp. 277–290.
30. Rizvi, B.; Belatreche, A.; Bouridane, A.; Watson, I. Detection of stock price manipulation using kernel based principal component analysis and multivariate density estimation. *IEEE Access* **2020**, *8*, 135989–136003. [CrossRef]
31. Leangarun, T.; Tangamchit, P.; Thajchayapong, S. Stock price manipulation detection using a computational neural network model. In Proceedings of the 2016 Eighth International Conference on Advanced Computational Intelligence (ICACI), Chiang Mai, Thailand, 14–16 February 2016; pp. 337–341.
32. Abbas, B.; Belatreche, A.; Bouridane, A. Stock price manipulation detection using empirical mode decomposition based kernel density estimation clustering method. In *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys)*; Springer: Cham, Switzerland, 2019; Volume 2, pp. 851–866.
33. Kakde, Y.; Chavan, G.; Sah, B.; Sen, A. Solution Approach for Detection of Stock Price Manipulation by Market Operators. In *Smart Technologies in Data Science and Communication: Proceedings of SMART-DSC 2022*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 281–288.
34. Zhao, W.; Alwidian, S.; Mahmoud, Q.H. Adversarial Training Methods for Deep Learning: A Systematic Review. *Algorithms* **2022**, *15*, 283. [CrossRef]
35. Shao, S. The effectiveness of supervised learning models in detection of pump and dump activity in Dogecoin. In Proceedings of the Second IYSF Academic Symposium on Artificial Intelligence and Computer Engineering, Xi'an, China, 8–10 October 2021; pp. 356–363.
36. La Morgia, M.; Mei, A.; Sassi, F.; Stefa, J. Pump and dumps in the bitcoin era: Real time detection of cryptocurrency market manipulations. In Proceedings of the 2020 29th International Conference on Computer Communications and Networks (ICCCN), Honolulu, HI, USA, 3–6 August 2020; pp. 1–9.
37. La Morgia, M.; Mei, A.; Sassi, F.; Stefa, J. The doge of wall street: Analysis and detection of pump and dump cryptocurrency manipulations. *ACM Trans. Internet Technol.* **2023**, *23*, 1–28. [CrossRef]
38. Chadalapaka, V.; Chang, K.; Mahajan, G.; Vasil, A. Crypto Pump and Dump via Deep Learning Techniques. *arXiv* **2022**, arXiv:2205.04646.
39. Victor, F.; Hagemann, T. Cryptocurrency pump and dump schemes: Quantification and detection. In Proceedings of the 2019 International Conference on Data Mining Workshops (ICDMW), Beijing, China, 8–11 November 2019; pp. 244–251.
40. Mirtaheri, M.; Abu-El-Haija, S.; Morstatter, F.; Ver Steeg, G.; Galstyan, A. Identifying and analyzing cryptocurrency manipulations in social media. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 607–617. [CrossRef]
41. Coin Market Cap. Available online: https://coinmarketcap.com/ (accessed on 23 June 2023).
42. Hu, S.; Zhang, Z.; Lu, S.; He, B.; Li, Z. Sequence-based target coin prediction for cryptocurrency pump-and-dump. *Proc. ACM Manag. Data* **2023**, *1*, 1–19. [CrossRef]
43. Chen, W.; Xu, Y.; Zheng, Z.; Zhou, Y.; Yang, J.E.; Bian, J. Detecting "Pump & Dump Schemes" on cryptocurrency market using an improved Apriori Algorithm. In Proceedings of the 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE), San Francisco, CA, USA, 4–9 April 2019; pp. 293–295.
44. Ccxt. Available online: https://github.com/ccxt/ccxt (accessed on 26 June 2023).
45. Mansourifar, H.; Chen, L.; Shi, W. Hybrid cryptocurrency pump and dump detection. *arXiv* **2020**, arXiv:2003.06551.
46. Bello, A.; Schneider, J.; Di Pietro, R. LLD: A Low Latency Detection Solution to Thwart Cryptocurrency Pump & Dumps. In Proceedings of the 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Dubai, United Arab Emirates, 1–5 May 2023.