Taxonomy of Fraud in Crowdfunding: A Hybrid Review of Detection Tools using Artificial Intelligence and Blockchain for Sustainable Platforms

loana Florina Coita[b,p,q,**], Marcos R. Machado[a,**], Lucia Gomez Teijeiro[d], Karsten Wenzlaff[h^j], Andreas Gregoriades[e], Christos Themistocleous[e], Frederik Sinan Bernard[a], Ramona Rupeika-Apoga[k], Huei-Wen Teng[l,m], Anastas Dzurovski[f], Gokce Nur Yilmaz[n], Liana Stanca[g], Wouter van Heeswijk[a], Marius Vlad Pop[b], Karolina Bolesta[c], Jana Peliova[p], Olivija Filipovska[o]

[a] *University of Twente, Department of Industrial Engineering and Business Information Systems, AE Enschede, 7500, NL*

[b] *University of Oradea, Faculty of Economics, Department of Finance and Accounting, Universitatii Str. 1,, Oradea, 410100, RO*

[c] *Warsaw School of Economics, Department of Economics I, Warsaw, 02-554, PL*

[d] *Bern University of Applied Sciences, Institute of Applied Data Science & Finance, Bruckenstrasse 73, Bern, 3005, CH*

[e] *Cyprus University of Technology, Faculty of Communication and Media Studies, 30 Arch. Kyprianos Str., Limassol, 3036, CY*

[f] *University St. Kliment Ohridski - Bitola, Kicevo Faculty of Law, st. Rudnichka b.b., Kicevo, 6250, MK*

[g] *Babes-Bolyai University, Faculty of Economic Science and Business Management, Department of Business Information Systems, Strada Teodor Mihali, Nr. 58-60, Cluj-Napoca, 400591, RO*

[h] *Hamburg University, Faculty of Business, Economics, and Social Sciences, Hamburg, 34396, DE*

[i] *University of Cambridge, Cambridge Centre for Alternative Finance, Cambridge, 1QA, UK*

[j] *University of Utrecht, European Centre for Alternative Finance, Utrecht, 3584CS, NL*

[k] *University of Latvia, Faculty of Economics and Social Sciences, Riga, LV-1050, LV*

[l] *National Yang Ming Chiao Tung University, Department of Information Management and Finance, Hsinchu, 300093, Taiwan, R.O.C.*

[m] *IDA Institute Digital Assets, Bucharest University of Economic Studies, Bucharest, 010374, RO*

[n] *TED University, Faculty of Engineering, Department of Computer Engineering, Ankara, 06420, TR*

[o] *Komercijalna Banka AD Skopje, Skopje, 1000, NMK*

[p] *University of Economics in Bratislava, Bratislava, 852 35, SK*

[q] *University of Sofia, St. Kliment Ohridski, Sofia, 1504, BG*

Abstract

Crowdfunding platforms have gained popularity as a means of financing entrepreneurial initiatives but face a high risk of fraud. Fraud erodes trust and causes financial instability. Thus, robust fraud detection and prevention are crucial for the sustainability of crowdfunding platforms. This study provides a systematic review of the literature and state-of-the-art discussions about crowdfunding fraud. Unsupervised topic modeling highlights that both AI and blockchain are recurrently presented in the literature as effective methodologies for identifying and preventing fraudulent practices.

Furthermore, this work describes current market practices of crowdfunding platforms in preventing fraudulent behavior and argues that while fraud is rare, its high impact requires new and innovative forms of fraud detection. A key limiting factor for the application of Artificial Intelligence (AI) solutions is the lack of available labeled crowdfunding data for training efficient algorithms for fraud detection, which is crucial as it constitutes an anomaly detection Machine Learning (ML) task. In this context, unsupervised ML methods are discussed as valuable techniques for detecting anomalies in the absence of labeled fraud cases because ofe of their ability to adapt to evolving fraud patterns.

This research provides valuable insights into the taxonomy of frauds and the complexity of detecting and preventing fraudulent activities in crowdfunding.

Keywords: Fraud Detection, Crowdfunding, Finance Industry, Alternative Finance Methods, AI guided Literature Review

1. Introduction

Financial fraud is a pervasive and evolving threat that undermines the integrity of financial systems worldwide. Deceptive practices are diverse and overall imply the intentional manipulation of information for personal gain. The rise of online finance platforms has increased the complexity of fraudulent strategies, requiring more sophisticated detection methods.

Among online finance platforms, crowdfunding platforms are in a widespread focus of interest, as they have introduced a modern way to finance entrepreneurial ventures by collecting funds from potential customers and investors (Teichmann et al., 2023; Lau et al., 2018; Pandey et al., 2019). Despite their popularity and expansive adoption, words of caution are numerous, and international institutions such as the Financial Action Task Force (FATF) have argued that crowdfunding might be used for fraud, terrorism financing and money laundering, as they presumably allow the international transfer of funds without any checks (Robock, 2014). However, crowdfunding platforms are under tight regulation, even in emerging and developing countries (Wenzlaff et al., 2021). In the last ten years, regulations have introduced anti-money laundering provisions, requiring platforms to identify customers. Campaign owners and supporters must provide details about their residence, tax status, and the intended use of funds. In addition, most crowdfunding platforms do not accept cash. Instead, they use digital payment providers that must comply with additional antimoney laundering regulations, such as flagging fraudulent transactions. However, fraudsters do exploit vulnerabilities of digital platforms, though rarely.

In this paper, as elaborated in section 3.1, fraud in crowdfunding platforms refers to any intentional act of deception or misrepresentation by campaign owners or participants to secure financial gain or other benefits under false pretenses. This includes, but is not limited to, the creation of fake campaigns, misrepresentation of project details, misuse of funds, and failure to deliver promised rewards or returns.

Fraudulent activities undermine the trust and integrity of crowdfunding platforms, posing significant risks to both investors and the platforms themselves. In this context and discussed further in Section 3.2, this paper makes an important distinction between crowdfunding platforms with versus without financial return, as the underlying motivations for fraud, incidence, and applied strategies might differ between that (Shneor et al., 2023). In the latter, so-called donation-based or reward-based crowdfunding platforms, after supporting a project, a supporter would receive a tangible reward or simply donate. Platforms of this type usually have a large number of crowdfunding campaigns and their owners have to submit information about the feasibility of the campaign to the platform as well as characteristics of the project or product (Wessel et al., 2022). If the crowdfunding campaign is successful, then crowdfunding platforms withhold the money until the identified checks of the campaign owner are finalized. Examples of fraudulent behavior on non-financial returns have been cited in the literature and in media articles covering the topic (Cardona et al., 2024; Cumming et al., 2021). The debate often centers around projects that have had a successful crowdfunding campaign but did not deliver the products that they advertised. This can be due to the intentional defrauding of supporters, or simply because campaign owners underestimated the efforts to deliver the products that they advertised (Rodriguez-Garnica et al., 2024).In contrast, this research paper distinguishes crowdfunding platforms with financial returns with crowdfunding platforms with non-financial returns, because of different risks of fraudulent behavior (Jin, 2024). Equity-based and lending-based crowdfunding platforms facilitate the transfer of money from the investor to the campaign and then from the campaign to the investor if the investment is successful. Therefore, these platforms are tightly regulated. For instance, in the European Union, the European Crowdfunding Service Provider Regime requires the platforms to verify the identity of investors - often by submitting a scan of the passport or using a digital identity tool before on-boarding the investor. In addition, platforms have to verify much more information about the investment project, for instance, the details of the incorporation of the business, the ownership structure, business plans and financial plans related to the investment project, criminal records of the people benefiting from the investment, and many other details about the project. Crowdfunding Platforms with financial returns are required to have in place mechanisms to detect fraudulent behavior on the platforms and report incidents to the regulatory authorities, especially when suspecting cases of money laundering (Wenzlaff et al., 2022). Therefore, the paper conceptualizes fraudulent behavior according to the different types of crowdfunding.

However, both types of crowdfunding platforms share a chronic common dilemma, which is a typical challenge for two-sided markets (Lacan & Desmet, 2017): Campaigns seek platforms that have a large number of users. At the same time, potential users are drawn to a platform by the quality of the campaigns. Fraudulent behavior on the platform undermines the reputation of the platform, thereby decreasing its attractiveness to new campaigns and new users. Academic literature has well established that the supporters trust the platform in the selection of projects with a high quality (Moysidou & Hausberg, 2020). Therefore, donation-based and reward-based crowdfunding platforms improve campaign quality by offering coaching.

Equity-based platforms rely on extended due diligence, while lending-based platforms often collaborate with loan originators, such as traditional banks, therefore providing a loan portfolio with a high quality (Schwartz, 2018).

This research focuses on the pressing issue of fraud in the domain of alternative financing (Lee et al., 2022; Lu et al., 2018; Chen & Wei, 2023), with a particular emphasis on crowdfunding platforms. The objective is to explore effective methodologies for the identification and prevention of fraudulent practices that are essential for sustaining the credibility and reliability of these innovative financial mechanisms.

The process of fraud detection in crowdfunding, as elaborated in section 3.3, involves the application of data-centric strategies to spot atypical behaviors indicative of fraud (Xu et al., 2015; Choi et al., 2022; Perez et al., 2022; Xu et al., 2022). This paper examines how patterns and trends are explored within extensive datasets in literature, such as transaction details, user interactions, and digital footprints, these platforms can identify discrepancies that stray from normal behavior. We have found that recent advancements in AI and predictive modeling have significantly bolstered the ability to detect fraud dynamically (Chandola et al., 2009; Kou et al., 2004). These technologies enable the ongoing surveillance and swift mitigation of suspicious activities, thereby minimizing the economic repercussions associated with fraudulent incidents. We have found evidence that implementing robust fraud detection protocols allows platforms to avert such fraudulent activities, ensuring equitable practices and compliance with regulatory frameworks.

To summarize shortly, this paper makes several significant contributions to the field of crowdfunding fraud detection. First, this paper provides a comprehensive and systematic literature review enhanced by employing an automated topic modeling approach. This dual-method approach ensures a thorough and unbiased review of the literature, synthesizing current knowledge on crowdfunding fraud. Second, the study categorizes various types of fraud strategies prevalent in crowdfunding platforms, offering a clear taxonomy that can be used by researchers and practitioners to understand better and identify fraudulent activities. This structured understanding of different fraud types fills a critical gap in the literature. Third, the paper evaluates the effectiveness of blockchain technology and ML models in detecting fraudulent activities. By examining these tools, the study provides insights into their strengths and limitations, guiding future research and practical applications in fraud detection. Fourth, the research addresses significant issues related to data availability and quality in crowdfunding applications. It discusses the challenges posed by limited labeled data and explores unsupervised methods such as clustering, autoencoders, and one-class SVMs as viable solutions for anomaly detection in the absence of labeled fraud cases. This discussion highlights potential solutions to data challenges in the field.

Here, derived insights aim to support both industry professionals, by providing a detailed understanding of advanced fraud detection techniques, and academic researchers, by laying the foundation for further inquiries into cutting-edge fraud prevention solutions in crowdfunding.

2. Methodology

To obtain a systematic overview of the state of current literature covering crowdfunding fraud, while ensuring spot-thematic recurrence, we implemented an SLR methodology combining a careful search-based pipeline (Figure 1) and Large Language Model (LLM)-based topic modeling (Figure 2). This approach combines a meticulous search-based pipeline, which rigorously identifies and filters relevant studies from diverse academic databases, with LLM-based topic modeling. The search-based pipeline ensures comprehensive coverage and inclusion of high-quality sources, while the LLM-powered analysis enables sophisticated extraction and clustering of recurring themes, key insights, and emerging trends. Together, these techniques provide a comprehensive framework for synthesizing the landscape of research on crowdfunding fraud, offering both quantitative and qualitative perspectives.

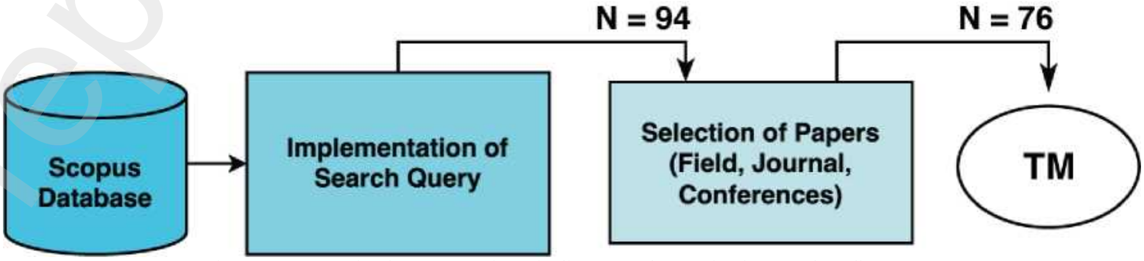

Scopus Database
Implementation of Search Query
N = 94
Selection of Papers (Field, Journal, Conferences)
N = 76
TM

Figure 1: Systematic Literature Review Pipeline - Fraud Detection in CrowdFunding.

Article collection was performed by keyword-based paper retrieval using theScopus[1] database. The authors not that the choice to focus on the latter database was twofold. Firstly, it allowed for a broader coverage of publications inclusive of not only journal but also relevant conferences and book chapters, and secondly, due to a vast coverage of recent crowdfunding, tech-related research. The search resulted in an initial pool of 94 papers related to fraud and crowdfunding fraud. Specifically, we used the search query: (“Fraud" OR “Fraudulent Activities" OR “Fraud Detection") AND (“Crowdfunding" OR “Alternative Financ*"). Retrieved papers were published between 1992 and September of 2024, though analysis was limited to those published within a timeframe of the last ten full years from 2013 onwards. This was to ensure relevance to contemporary discussions and modern technology applications. The initial pool of retrieved papers followed a human-based careful selection process consisting of three sequential quality control phases. Firstly, we filtered the papers by field, journal, and/or conference publishing, process resulting in a reduced set of 76 papers. To refine the initial pool, exclusion criteria were applied to ensure higher relevance and focus. Specifically, papers not published in English were excluded, resulting in the removal of non-English entries such as Chinese and Korean papers. Regarding document type, only journal articles, conference papers, and reviews were retained, leading to the exclusion of book chapters and other document types. Subject area filters were also applied, prioritizing fields directly relevant to fraud detection and crowdfunding, such as Computer Science, Engineering, Mathematics, and Decision Sciences. Papers categorized under unrelated fields such as Environmental Science, Chemistry, and Medicine were excluded from further consideration.

Secondly, we leveraged state-of-the-art AI Topic Modeling both to define the thematic areas to be discussed (see Figure 2 and Main Findings section) and to further prune peripheral thematic areas for which literature does not provide robust coverage. Thirdly, authors carefully examined papers’ titles, keywords, abstracts and topical assignment towards a final set of 76 research works.

To guide the literature review, as a visualization technique, we have applied unsupervised ML on papers’ titles, abstracts, and sets of highlighted keywords both by the author and Scopus indexing system. To do so, we preprocessed the text for format homogeneity using the topmost Python module (lowercasing, stop-word removal) (Wu et al., 2023b). Next, we applied state-of-the-art topic modeling using the Fastopic python module, published in 2024 by NeurIPS and declared as the most stable and performant transformer-based model given the implementation of Dual Semantic-relation Reconstruction technology (Wu et al., 2023a). We customized its implementation to suit literature topic modeling by calling the Allenai specter transformer model, callable through the sentence-transformers API, and by setting the preprocessing vocabulary size to 2852 tokens, the corpus size of our paper database. We tested several topics in training experiments of 1000 epochs and chose the optimal number of topics by inspecting the percentage of topic diversity captured. With five topics we reached a 94.6%.

We applied a standard R pipeline for Latent Dirichlet Allocation (LDA) Topic Modeling (3). LDA results were used for defining the results subsections covered, as summarized in Figure2: fraud definitions and types, blockchain methods for crowdfunding fraud detection and prevention, ML methods for crowdfunding fraud detection and prevention, and crowdfunding use cases for which fraud constitutes a red flag problem.

[1] www.scopus.com

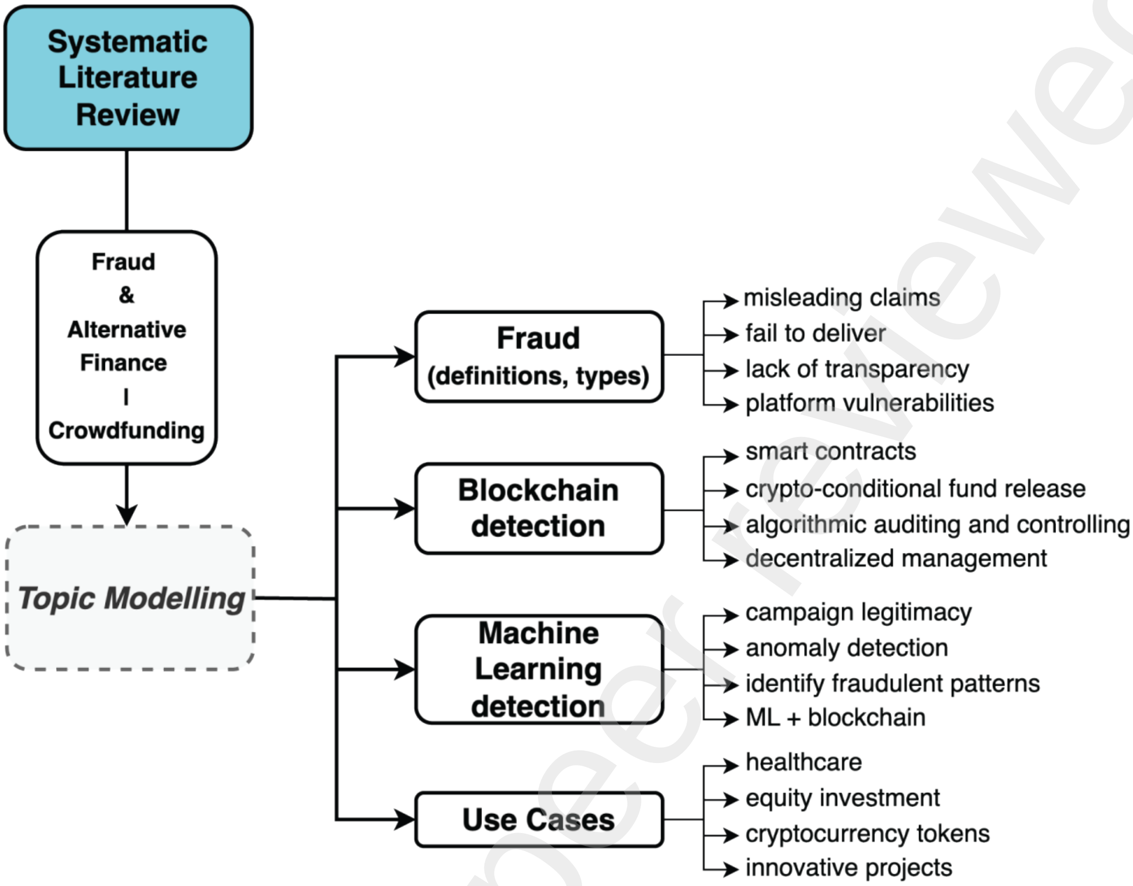


Figure 2: Topic Modeling output.

LDA used to discover the underlying topics in a collection of text documents is based on the assumption that each document is composed of a mixture of topics and that each topic is a distribution over a fixed vocabulary of words (explicitly present in the documents). One-word engrams were provided as input for LDA, embedded using their TF-IDF (term frequency-inverse document frequency) weights, which measure the importance of words in each document relative to the entire corpus. LDA results indicated that the 54 analyzed papers stably align to five overarching topics: 1) platform factors in crowdfunding fraud, 2) crowdfunding and blockchain, 3) models and platforms for crowdfunding fraud, 4) crowdfunding for healthcare, and 5) crowdfunding for equity investment (Figure 3). These topic titles were derived from LDA beta weights, while paper alignment to topics was retrieved from LDA alpha weights. Constituting a suitable topic modeling method due to its high explainability and easiness of interpretation, LDA is however sensitive to the corpus size and hyperparameter tuning, thus a word of caution relative to the limited scope of available literature must be made.

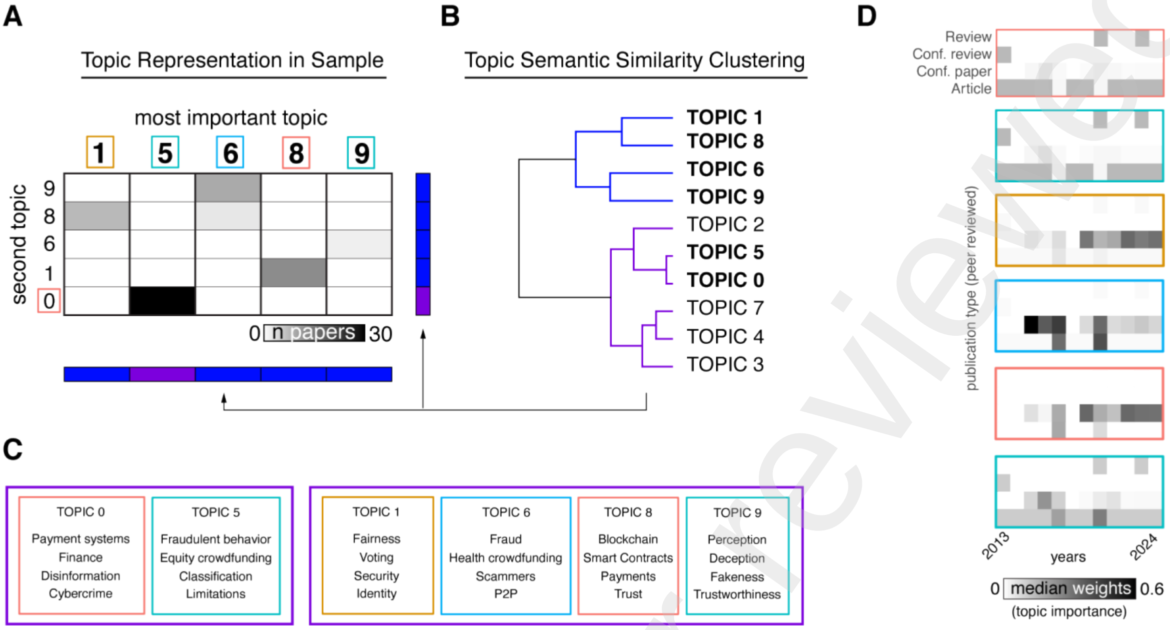


Figure 3: Topic Modeling Results. Representative topics and time evolution.

To further provide a glance view of the covered literature and the topical and authorship relations, we used ResearchRabbit[2], for constructing a Network of this paper-collection (Figure 4). Results showed that the expert-selected papers cover the diversity of approaches for Fraud in Crowdfunding, AI and blockchain methods for its detection, diverse case studies across sectors, and potential solutions using AI and Blockchain implementations to crowdfunding platforms. We also represented this network of papers across their timeline of publication (Figure 4), seeing that, in concordance to what was discovered through Scopus[3] retrieval, the number of publications on this matter experiences and exponential growth over time, confirming its relevance. Results from topic modelling also highlighted similar topics thus providing evidence that such approaches can be used in combination.

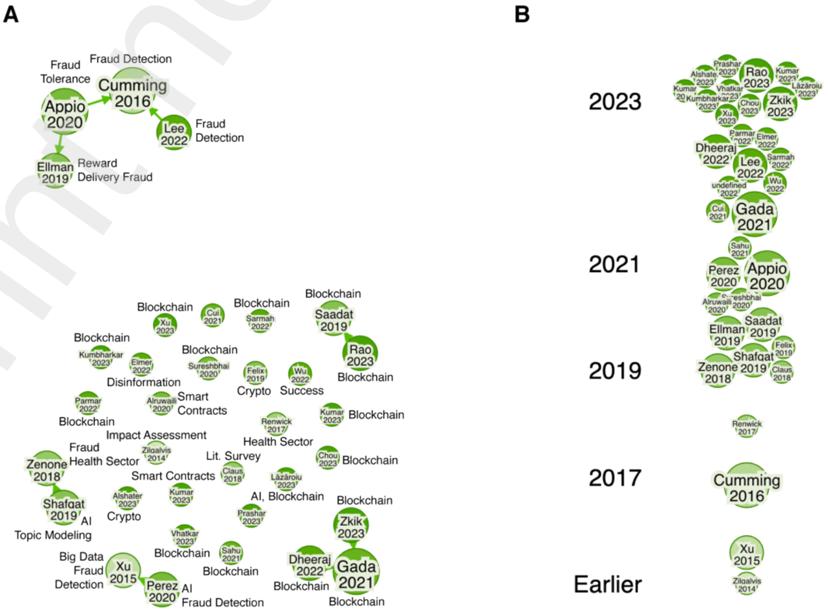


Figure 4: Network representation of Crowdfunding literature, topics, and temporal evolution. The subset of analyzed papers - ResearchRabbit;

3. Main findings

3.1. Definitions of Fraud

Crowdfunding fraud can occur when campaign creators (intentionally) deceive backers to secure funds without intending or being able to fulfill promised rewards or project goals. In reward-based and donationbased crowdfunding, it involves presenting false information about the product or overstating feasibility, leading backers to support projects based on misleading claims Cumming et al. (2021). Using compelling pitches to attract funding but then fail to deliver or disappear after receiving funds is sometimes considered fraud in the academic literature, even though fraudulent intentions might not exist (Cumming et al., 2021; Wang, 2019). Since donation-based and reward-based crowdfunding platforms verify less rigorous than equity-based and lending-based platforms, reward-based backers are more vulnerable to fraud, making transparency and platform accountability critical concerns (Liu et al., 2023). Mbarek & Trabelsi (2020) characterize fraud as an intentional scheme by a party that deliberately seeks to create a false impression to gain an unfair advantage over another party.

Wang & Wang (2019) defines internet fraud as the deliberate manipulation of charitable (donation-based) crowdfunding platforms by agents, who create multiple fake projects to deceive donors and illicitly acquire funds. In a different context, Akhmadiyev et al. (2023a) describes pyramid scheme fraud as a structure where income is generated not through actual economic activity but by attracting new participants, whose funds are used to pay earlier participants. These definitions highlight the reliance of fraudulent activities on trust, information asymmetry, and the necessity for stringent verification mechanisms in online and financial platforms.

Fraud in Initial Coin Offerings (ICOs), which some scholars consider a form of crowdfunding without specific platforms acting as intermediaries, often arises from the inherent lack of regulation and information asymmetry between project creators and investors. Kumar et al. (2023b) highlight the potential of using blockchain technology to prevent scams in crowdfunding platforms, where investors do not receive the promised rewards or products after making contributions. In the context of ICOs, Chou et al. (2023) discusses fraud as a significant issue caused by the lack of mandatory disclosure requirements due to missing intermediaries, which allows project creators to manipulate information in white papers without platforms verifying the information in the white papers. Furthermore, Siering et al. (2016a) provides a comprehensive definition of fraud in crowdfunding as deceptive behavior by project creators, focusing on the exploitation of linguistic and content-based cues to mislead investors.

3.2. Fraud Strategies

Fraud strategies in crowdfunding exploit inherent vulnerabilities within platforms capitalising on the behavior of potential contributors.

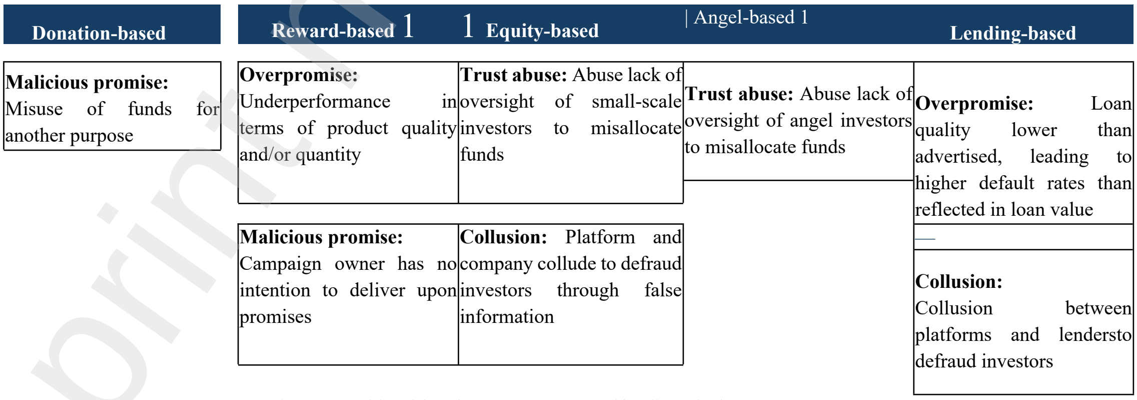
Prolonging the funding period in reward-based crowdfunding, campaigns aim to maximize the amount of capital raised, as they lack credible signals to assure backers of their project's legitimacy (Cumming et al., 2021). Additionally, they create multiple pledge categories with low funding thresholds, which allows them to attract numerous contributors.

Another strategy involves the use of vague or misleading campaign descriptions. Fraudulent campaigns present unclear project details, making it challenging for contributors to assess the viability and legitimacy of a project (Siering et al., 2016b). This obfuscation can lead to confusion among potential backers. Furthermore, fraudsters may specifically target less-educated audiences, capitalizing on their lack of awareness regarding potential fraud indicators (Mbarek & Trabelsi, 2020). These tactics underscore the critical need for vigilance among contributors, as well as the importance of implementing robust mechanisms for fraud detection and prevention within crowdfunding platforms to protect the supporters. In the domain of donation-based crowdfunding, Wang & Wang (2019) highlights how agents exploit the system by setting up multiple projects, thus gaining access to a wide base of donations without delivering the promised services. The primary strategy here is leveraging the platforms' limited capacity to verify the authenticity of each project effectively. In pyramid schemes, Akhmadiyev et al. (2023a) identify a distinct strategy where the scheme operators promise high returns to initial investors by using the funds from subsequent participants. The system collapses when new participants can no longer sustain the payments, leading to significant financial loss. The fraud strategy in pyramid schemes relies on attracting a continuous flow of new participants, exploiting their lack of financial awareness, and promising high returns without substantial risk.

Several fraud strategies are commonly employed in ICOs. Kumar et al. (2023b) describe how the publishers of white papers for ICOs use the collected funds for personal use rather than for the intended project development. They also point out the high transaction fees in traditional payment systems, which incentivize fraudulent behavior. Chou et al. (2023) emphasize the role of incomplete or misleading information in ICO white papers as a key strategy for fraud, where low-quality projects mimic the signals of high-quality projects to deceive investors. Siering et al. (2016a) identify specific linguistic and content-based strategies, such as exaggerated promises and manipulative language, used by fraudulent ICOs to increase the perceived legitimacy of their projects. They point out that static information about the project is likely to be intentionally falsified by deceivers.

3.3. *Fraud Varies across Crowdfunding Platform Types - Fraud types*

Various types of fraud can be identified across distinct types of crowdfunding platforms. Figure 5 summarizes the main types and is followed by an in-depth discussion per fraud type.



Donation-based	Reward-based 1	1 Equity-based	| Angel-based 1	Lending-based
Malicious promise: Misuse of funds for another purpose	**Overpromise:** Underperformance in terms of product quality and/or quantity	**Trust abuse:** Abuse lack of oversight of small-scale investors to misallocate funds	**Trust abuse:** Abuse lack of oversight of angel investors to misallocate funds	**Overpromise:** Loan quality lower than advertised, leading to higher default rates than reflected in loan value
	Malicious promise: Campaign owner has no intention to deliver upon promises	**Collusion:** Platform and company collude to defraud investors through false information		**Collusion:** Collusion between platforms and lendersto defraud investors

Figure 5: Table of fraud types across crowdfunding platform types

In donation-based crowdfunding, where no financial return is given, fraudulent behavior can occur if a campaign aims to contribute to a certain beneficial cause but then misuses the funds for another purpose. For instance, a non-profit organization might claim that funds might be used for combating climate change, but instead, the funds are used to pay for other expenses. Most countries oblige non-profit organizations to report their income (especially donations) and disclose their spending. Voluntary certificates of transparency are used to show that funds are used following the intention of the donors. Non-profit associations are often required to be audited to achieve these certificates. Donation-based crowdfunding platforms thereby rely on these certificates, auditing procedures and transparency requirements when on-boarding non-profits to the platforms (Salido-Andres et al., 2022).

Reward-based crowdfunding works on the premise that the final product will be delivered to financial backers after a specific development period and once the full pledged amount has been received (all-or-nothing approach). Unlike in equity-based or lending-based crowdfunding, as well as in traditional funding methods, such as venture capital or bank loans, companies using crowdfunding are not required to disclose details about their financial background or stability for public assessment. As such, prospective supporters make decisions based on descriptive information, such as the concept of the product, its detailed explanation, and the level of support from other backers. In addition, reward-based crowdfunding features entrepreneurs who might not have a long history of financial statements, as they are often start-ups or small enterprises (Se- waid et al., 2021). Fraudulent behavior in reward-based crowdfunding can take the following forms: a) The campaign owner might not provide the quality or quantity of the product as described. The campaign owner might be overwhelmed by the response to the crowdfunding campaign, and therefore underestimate the efforts to fulfill the campaign's promises. This may be unintentional; therefore, it would not be considered fraud in the strictest legal terms, but negligence on behalf of the campaign owners; b) The campaign owner might not have the intention to fulfill any of the promises of the campaign, simply proposing a campaign that is not feasible or realistic. In this case, this could be considered fraudulent behavior in the strict legal term.

While reward-based crowdfunding platforms usually exclude responsibility for the fulfillment of the campaign van Otterloo (2022), they have put in place measures to prevent this from happening. For instance, platforms such as Kickstarter require campaigns to submit evidence of their capacity to fulfill the campaign promises. They will also keep the pledged amount in an escrow account until certain milestones are reached by the campaign owner.

In equity-based crowdfunding, fraudulent behavior can be related to insufficient oversight of the investors over the company in which they have invested. As equity-based crowdfunding allows retail investors to invest small amounts in enterprises, especially in start-ups, they might not have the incentive nor the capability to exercise regular control over the activities of the company that has been funded. For instance, a start-up might use a crowdfunding campaign to collect funds for the expansion of the business, but then use the collected funds to spend on the salary of the CEOs (Rosli & Shahida, 2019).

This phenomenon is not unique to equity-based crowdfunding, it is a potential threat to any angel funding for start-ups (Van Osnabrugge, 2000). Typically business angels, which invest more than 25.000 EUR per ticket, address this problem by sharing oversight among each other, for instance by installing a lead investor, who has a seat on the Board of Directors of the start-up and is mandated to use the voting power of the shares of the other angels to ensure that the start-up is using the investments wisely and in accordance with the business plan. Investment contracts of business angels typically ensure that the owners of a start-up have the same incentives as their investors, for instance in maintaining the value of the intellectual property of the start-up.

Equity-based crowdfunding platforms have reacted in similar ways. For instance, most equity-based crowdfunding platforms collect the funds in Special Purpose Vehicles (SPV), which then invests in the start-up. The special purpose vehicle pools the investments of the retail investors, thus also pooling the voting power. In this particular case, the platform usually manages the SPVs, given that the platform has extensive knowledge of the business plans and financial plans of the start-ups (Hooghiemstra, 2022).

If the equity-based platform and the company seeking the investment collaborate with the intent to defraud, then fraudulent behavior might happen on an equity-based crowdfunding platform. However, in most equity-based crowdfunding regulations, some provisions prohibit the collusion of platform and project, for instance, the platform is not allowed to have an equity stake in the project seeking the financing (Duarte, 2022).

In lending-based crowdfunding, the platform typically intermediates loans, which are described based on the loan characteristics, such as maturity, interest rates, and risk category (Ziegler et al., 2021). The lender is not described in detail, other than maybe the name and category of the beneficiary of the loan. Usually, there are no campaign pages in the classical sense of crowdfunding.

Investors on lending-based crowdfunding platforms build their loan portfolios by selecting loans that match their risk preferences. Most lending-based platforms have now resorted to offering automatic portfolio investments. The lender indicates a risk preference and a maximum investment budget, the platform then assigns the loans to the lender based on this risk assessment (Ferretti, 2022).

Fraudulent behavior on lending-based crowdfunding platforms would necessitate significant collusion with criminal intent between the platform and several thousand lenders, attempting to collect the funds and then close the platform, for instance. Fraudulent behavior could also be done by platforms operating Ponzi schemes, whereby the interest rates of earlier investors are paid using investments from later investors. It should be noted that in the European lending market, this kind of criminal behavior has not been observed, but it has been the cause of very strong market regulation for lending-based crowdfunding in China (Huang & Pontell, 2023).

Another way in which investors might be damaged by lending-based crowdfunding would be if the loan quality is substantially inferior as claimed by the platform, thus leading to a higher rate of loan default than originally advertised. This has been observed in the European lending market, especially during the pandemic in 2021 and 2022 (Olvedi, 2022). The platforms suffered damage to their reputation, because consequentially investors retreated from the platforms, and some platforms went insolvent in the following years. However, this type of negligence on behalf of the platform is not uncommon in other markets, where due to external shocks the portfolio value has been reduced significantly. Lending-based crowdfunding platforms combat this phenomenon by being transparent about the method of calculating the risk category and by providing historical data on loan defaults. This is also required by the European Crowdfunding Service Provider Regime, as well as national and international requirements (Ferretti, 2022).

3.4. Fraud Detection through Blockchain

Several studies are researching blockchain technology as a valuable option against crowdfunding fraud. There is wide consensus among researchers that smart contracts (also known as crypto contracts) are the most viable option for utilization of the blockchain technology against crowdfunding fraud.

A Smart Contract can be defined as a program that directly and automatically controls the transfer of digital assets between the parties and verifies that certain conditions will be met (De Caria, 2020). There are many similarities between traditional contracts and smart contracts and the second is automatically enforcing the contract. Traditional contracts are enforceable by law while smart contracts are enforceable by code. Smart contracts execute exactly as they are coded.

Blockchain technology is increasingly recognized as an innovative solution for mitigating fraud in crowdfunding. Among its applications, smart contracts - or crypto-contracts - are emerging as the most viable option. Smart contracts are automated programs that directly facilitate the transfer of digital assets, automatically executing when predefined conditions are met. Unlike traditional contracts, which are enforceable by law, smart contracts are enforceable by code, ensuring accuracy and minimizing execution errors (Liu et al., 2023). Recent research has demonstrated various applications of blockchain in fraud prevention. Naik & Oza (2023) combine blockchain transparency with ML algorithms to detect and prevent fraud, while Kumar et al. (2023a) use smart contracts to release funds only when certain conditions are met, thus ensuring that campaign creators deliver on their promises. Sahu et al. (2021) highlight the use of blockchain transparency and smart contracts to prevent fraud, focusing on automating the disbursement of funds under strict conditions. Xu et al. (2023) introduce a broader blockchain-based trust management mechanism, which includes features such as an auditor committee selection algorithm, incentives for auditors, and detailed workflows for crowdfunding trust management. Meanwhile, Rajarajeswari et al. (2023) and Prashar & Gupta (2024) emphasize the role of decentralized ledgers in maintaining immutable records of transactions, further enhancing the security and reliability of crowdfunding platforms. Beyond crowdfunding, blockchain technology shows promise in various other domains. Wu et al. (2022) propose a fuzzy q-rung orthopair decision-making model for evaluating crowdfunding platforms in microgrid investments, enhancing reliability through sensitivity analysis. Lazaroiu et al. (2023) highlight the integration of AI, blockchain, and big data in fintech, promoting risk assessment and sustainability. Similarly, Fang & Stone (2021) demonstrates the role of blockchain in improving transparency and efficiency in the dairy supply chain through real-time IoT data and smart contracts. Rajarajeswari et al. (2023) are employing blockchain's transparency and security (through an Open Permissioned Blockchain Solution for Private Equity Funding Using a Global, Cross-Cloud Network Blockchain Platform) to build investor confidence and ensure the integrity of transactions. Cryptocontracts (also known as smart contracts) are frequently researched techniques for crowdfunding fraud detection and prevention. Naik & Oza (2023) are employing a combination of blockchain's transparency features and ML algorithms to detect fraud. Smart contracts are used to automate the release of funds only when predefined conditions are met. Sahu et al. (2021) are employing blockchain's transparency and crypto contracts to detect and prevent fraud. Smart contracts automate the release of funds only when predefined conditions are met, ensuring that campaign creators adhere to their promises. Liu et al. (2023) introduce a blockchain-based trust management mechanism for crowdfunding, 2) design an auditor committee selection algorithm, 3) implement incentives for auditors, 4) use blockchain technology and smart contracts for transparency and security, 5) detail the workflow for various processes in crowdfunding trust management.

Naik & Oza (2023), Sun et al. (2023) and Li et al. (2024) are highlighting the potential of decentralized ledgers in maintaining an immutable record of transactions. (Kumar et al., 2023a) are seeing integrating smart contracts and Blockchain technology into the prevalent crowdfunding process schemes as a key element for fraud prevention within crowdfunding.

Furthermore, studies in energy and FinTech explore innovative technologies and methods to improve decision-making and operational efficiency. (Wu et al., 2022) propose a q-rung ortho-pair fuzzy decision-making model for evaluating crowdfunding platforms in microgrid investments, enhancing reliability through sensitivity analysis. The two-stage approach offers a comprehensive analysis, contributing to energy management systems. (Lazaroiu et al., 2023) highlight the integration of AI, blockchain, and big data in Fin- tech, enhancing risk assessment and promoting sustainability. Similarly, Fang & Stone (2021) proposes a blockchain-based dairy supply chain solution, improving transparency, security, and efficiency with realtime IoT data and smart contracts. Such studies underscore the importance of advanced technologies in optimization.

Jadhav et al. (2023) are employing achievement of validation through the consensus mechanism of the Ethereum blockchain, specifically using Proof of Virtual Voting (POVV) for verifying transactions.

Blockchain's distributed ledger system can provide transparent, immutable records of transactions, potentially addressing issues such as the fraudulent creation of multiple projects by agents, as discussed by Wang & Wang (2019). This approach could enhance trust by ensuring that every donation and transaction is permanently recorded, and any manipulation of project outcomes or funds can be audited transparently. Similarly, Naik & Oza (2023) and Sun et al. (2023); Li et al. (2024) highlight the potential of decentralized ledgers in maintaining an immutable record of transactions.

Blockchain technology has the potential to address fraud in ICOs by increasing transparency and providing a decentral alternative to intermediaries. Kumar et al. (2023b) propose a blockchain-based payment system utilizing Ethereum smart contracts to create a scam-proof arrangement between investors and project creators. This system ensures that funds are only released when specific conditions are met, and verified by the community of investors, thereby reducing the likelihood of fraud. In the context of ICOs, Chou et al. (2023) suggests that the transparency inherent in blockchain technology, combined with regulated security token offerings (STOs), can improve investor trust and reduce the potential for fraudulent behavior.

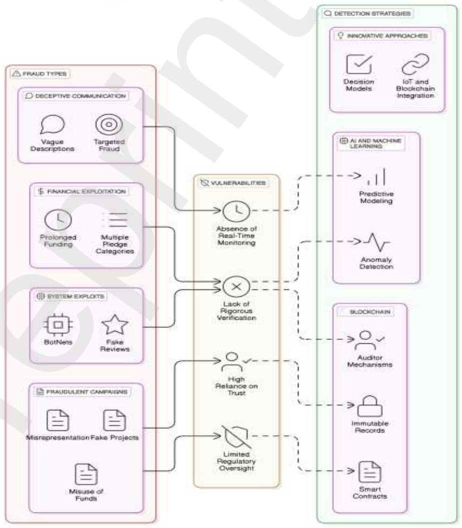


Figure 6: Types of Fraud (Visualization based on the findings of this study)

3.5. Fraud Detection through Machine Learning

Traditional rule-based systems identify potential fraudulent activities by relying on established rules and patterns. However, these systems find it challenging to adapt to emerging fraudulent behaviors, leading to numerous false negatives and possible financial losses. Consequently, there has been significant interest in utilizing ML algorithms to overcome these limitations. Hence, ML models mark a new era in detecting potential fraud. A well-structured methodology is aimed to be presented by recent ML research to expand current knowledge on fraud detection Hernandez Aros et al. (2024). Natural Language Processing (NLP) has become a prevalent way to detect potential fraud. In the context of campaign legitimacy, Perez et al. (2022) identified the language of legitimate campaigners on GoFundMe to be more descriptive and informative compared to fraudsters, a finding that was complemented by Cumming et al. (2023) results of language complexity in Kickstarter campaign descriptions being associated with legitimacy.

In addition to language processing, ML is capable of identifying anomalies from other project features. Wang & Wang (2019) discuss patterns of agent activity in donation-based crowdfunding platforms, such as creating numerous projects, which could be identified through anomaly detection algorithms. Similarly, pyramid schemes could be flagged by identifying abnormal transaction patterns and clustering techniques, as Akhmadiyev et al. (2023a) suggest, though no ML model was applied in their work. Supervised learning models, including decision trees and random forests, could help identify fraudulent patterns from large datasets, while unsupervised techniques like clustering could highlight suspicious activities in scenarios with limited labeled data.

The integration of ML techniques offers powerful tools for detecting fraudulent behavior in online platforms. Siering et al. (2016a) apply ML-based text mining techniques to analyze both the linguistic and contentbased cues of crowdfunding projects. By extracting features such as sentiment, readability, and writing style, their approach effectively distinguishes between legitimate and fraudulent projects. Choi et al. (2022) employs a hybrid ML approach that combines NLP and collaborative filtering, applied to detect fraudulent activity for a GoFundMe project portfolio in the healthcare sector. The experimental results demonstrate the added value of integrating several ML algorithms, a finding that is echoed by the study of Lee et al. (2022) that employs forward stepwise logistic regression and analyzes both linguistic and content-based cues.

An emerging branch of ML applications for fraud detection in crowdfunding ties to blockchain-based platforms. Although Kumar et al. (2023b) and Chou et al. (2023) do not explicitly implement ML techniques, their discussions suggest that integrating blockchain-based systems with ML could further enhance fraud detection by identifying anomalous patterns in transaction data and white paper content. In a similar vein, Zkik et al. (2024) employs Graph Neural Networks to detect and prevent smart contracts-based attacks in blockchain-based crowdfunding platforms.

3.6. Issues with Data Sources in Crowdfunding Applications

The literature on generating or synthesizing class labels (whether it is fraud or another type of anomaly) is extremely scarce. The general approach in this line of research is to apply unsupervised learning (such as K-means) on unlabelled data to predict labels, then use these labels to train a supervised learning model, and finally compare the performance of this model to a supervised learning model on actual labels. This approach manifests itself in works such as Baek et al. (2021); Moslehi et al. (2020); Maqbool & Babri (2006); Rauber (1999); Kennedy et al. (2024).

Baek et al. (2021) applied K-means clustering to estimate binary labels for cyber-network anomalies based solely on features and then used a supervised model with these labels to classify networks as anomalous and non-anomalous. According to their results, the supervised model with estimated labels performed very closely to the model with original labels. Moslehi et al. (2020) proposed an approach for assigning labels to clusters in a dataset. They use a labeled data set along with K-means clustering to improve the labeling of another, unlabelled dataset. In a leading work, Kennedy et al. (2024) tackled the challenges of imbalanced and unlabelled credit fraud data. They used an auto-encoder that learns from unlabelled data in an unsupervised manner to calculate an error metric, which was then used to synthesize binary class labels.

Additionally, unsupervised methods such as clustering can be applied to identify outliers in the absence of labeled fraud cases. Autoencoders and one-class SVMs are useful techniques when fraudulent labels are sparse or unavailable, providing anomaly detection capabilities that can adapt to evolving fraud patterns. The fraud patterns highlighted by Akhmadiyev et al. (2023a) in pyramid schemes could be well-suited for such approaches, as pyramid schemes often involve subtle deviations from legitimate financial behavior.

ML techniques offer powerful tools for detecting fraudulent behavior in online platforms. Siering et al. (2016a) apply ML-based text mining techniques to analyze the linguistic and content-based cues of crowdfunding projects. By extracting features such as sentiment, readability, and writing style, their approach effectively distinguishes between legitimate and fraudulent projects. Although Kumar et al. (2023b) and Chou et al. (2023) do not explicitly implement ML techniques, their discussions suggest that integrating blockchain-based systems with ML could further enhance fraud detection by identifying anomalous patterns in transaction data and white paper content.

3.7. Addressing the Data Imbalance Challenge in Fraud Detection

Fraud detection often deals with imbalanced and unlabeled data, as most transactions are legitimate, and fraudulent activities are rare. Neither Wang & Wang (2019) nor Akhmadiyev et al. (2023a) address this challenge directly, but their studies suggest areas where such data issues arise. In the case of donationbased crowdfunding, a large dataset of donation transactions may have only a few fraudulent instances, which would lead to class imbalance.

Data imbalance often causes models to favor predictions for the majority class, leading to the underrepresentation of the minority class and degrading overall model performance (Chen et al., 2024). To build robust ML models, addressing this imbalance is crucial. Techniques such as resampling, class weighting, and employing more suitable evaluation metrics play a vital role in enhancing model performance when working with imbalanced datasets.

Resampling techniques modify the dataset by either increasing the minority class samples (oversampling) or reducing the majority class samples (undersampling) (Moreo et al., 2016; Liu & Tsoumakas, 2020). A widely used method is the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic minority samples by interpolating between existing instances. This approach helps balance the dataset and enhances the model's ability to generalize to the minority class (Chawla et al., 2002). Drummond et al. (2003) compare various sampling techniques and emphasize the trade-offs between undersampling and oversampling.

Many ML algorithms allow for assigning higher weights to the minority class, penalizing misclassifications more heavily and encouraging the model to focus on improving predictions for the minority class. King & Zeng (2001) propose weighting schemes that modify the loss function to account for rare events in logistic regression, which can also be applied to other classifiers. Researchers have explored cost-sensitive approaches, such as cost-sensitive decision trees (Sahin et al., 2013) and cost-sensitive neural networks (Yotsawat et al., 2021). Additionally, imbalance-aware loss functions like Focal Loss (Lin et al., 2017) and Dice Loss (Li et al., 2019) have been developed. While these techniques help address data imbalance, they still face challenges such as overfitting, information loss, and algorithm-specific limitations.

In the context of imbalanced data, traditional evaluation metrics like accuracy can be misleading, as they may not reflect a model's true performance across both majority and minority classes. Alternative metrics such as precision, recall, and the F1-score provide a more accurate assessment of the model's effectiveness on imbalanced datasets. Jeni et al. (2013) recommend using metrics like the Area Under the Curve (AUC) and the F1-score to evaluate classifiers in these settings more reliably.

3.8. Application Cases in the Literature

As a result of the topic modeling implementation described in Section 2, one of the resulting topics reveals a cluster of literature focused on the intersection of crowdfunding and healthcare, and another on crowdfunding and blockchain.

Medical crowdfunding is an important financial tool in medical sector, it allows patients and health organizations to address financial barriers that might otherwise limit access to necessary care (Wenzlaff, 2023). Healthcare crowdfunding campaigns often seek to cover medical expenses, research funding, and community-based health initiatives. This section explores healthcare crowdfunding through specific case studies, highlighting the growing relevance of this funding model.

In the context of healthcare crowdfunding, Renwick & Mossialos (2017) discusses how patients with chronic conditions like diabetes in the United States are increasingly turning to crowdfunding platforms such as GoFundMe to cover medical and associated costs. Many diabetes patients face significant financial hardships, even when insured, and resort to crowdfunding to cover expenses beyond direct medical care, such as transportation, healthy food, and diabetic alert dogs. The analysis reveals that only 14% of crowdfunding campaigns reach their financial goals, suggesting that while crowdfunding offers a potential lifeline, it is often insufficient in addressing the entire financial burden that healthcare imposes on patients. The study also highlights indirect expenses as significant contributors to financial stress, which underscores the limitations of both healthcare policies and crowdfunding as sustainable financial solutions in the healthcare system (Renwick & Mossialos, 2017).

Sloan et al. (2023) further elaborates on the role of crowdfunding in healthcare by providing a typology of health-related crowdfunding projects. These include campaigns aimed at covering individual health expenses, funding health-related research, and financing commercial health innovations. While crowdfunding democratizes access to funding and raises awareness for overlooked health issues, Sloan et al. (2023) points out significant risks, such as inefficient priority setting, fraud, and regulatory gaps, which can hinder the broader goal of public health equity. The economic structure of crowdfunding health campaigns, according to Sloan et al. (2023), brings both opportunities for increased market participation and threats of market failure due to moral hazard and adverse selection, where financial aid may be misallocated (Sloan et al., 2023).

On the other topic, crowdfunding, when integrated with blockchain technology, offers a novel approach to financing projects, especially in emerging sectors like cryptocurrency and equity investments. The decentralized and transparent nature of blockchain aligns well with the crowdfunding model, introducing new possibilities but also new risks.

Felix & von Eije (2019) investigates underpricing in Initial Coin Offerings (ICOs), a form of blockchainbased crowdfunding used to raise capital by offering cryptocurrency tokens to investors. The study demonstrates that ICOs experience significantly higher levels of underpricing compared to traditional Initial Public Offerings (IPOs). The research reveals that U.S.-based ICOs, in particular, showed an average underpricing of 123%, which is even higher than IPO underpricing during the dot-com bubble. ICOs, like other forms of crowdfunding, are characterized by asymmetric information, which can lead to significant market volatility. Felix & von Eije (2019) highlights how factors such as first-day trading volume and positive market sentiment exacerbate the levels of underpricing, benefiting early investors but potentially reducing long-term gains for issuers. The paper also draws attention to the regulatory challenges ICOs face, suggesting that improved data transparency and stricter regulations could help reduce fraud and information asymmetry in the blockchain crowdfunding market (Felix & von Eije, 2019).

In contrast, Yeon et al. (2022) explores the legal implications and risks associated with equity crowdfunding (ECF), especially when combined with blockchain platforms. It examines how ECF enables startups to raise capital by offering small equity shares to investors through online portals but also discusses the vulnerabilities to cybercrime and fraud that arise from this digital platform. The article critically assesses the legal frameworks in Malaysia, such as the Capital Market and Services Act 2007 and the Securities Commission's Guidelines on Recognized Markets 2020, which aim to regulate equity crowdfunding and protect against cyber threats. Yeon et al. (2022) finds that while regulations exist, they leave gaps in addressing issues like intellectual property theft and compliance with public offering rules. It argues that more robust legal protections are needed to safeguard both issuers and investors in the blockchain-based crowdfunding space (Yeon et al., 2022).

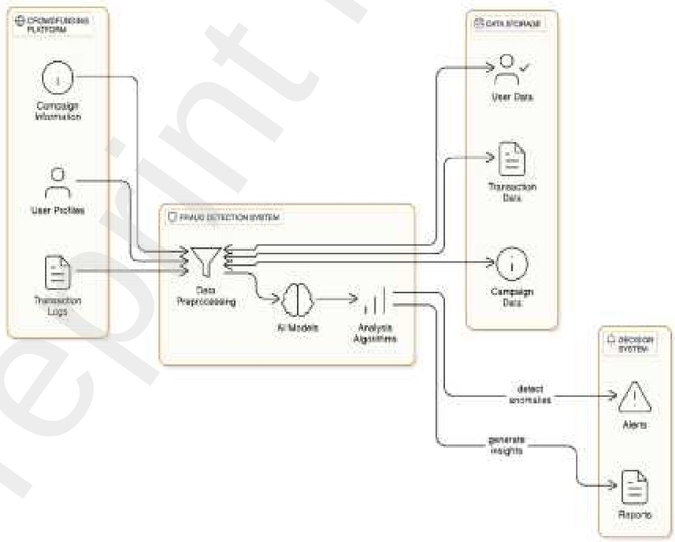


Figure 7: Crowdfunding Fraud Detecting Architecture (Visualization based on the findings of this study)

4. Final Considerations

In the realm of global finance operations, fraud has emerged as a significant problem, and crowdfunding platforms are not immune to the problems that it causes. The limited availability of labeled data is one of the key obstacles that must be overcome to detect fraudulent activity within the realm of crowdfunding. Because of this paucity, the effective application of supervised ML methods is hindered. These approaches, which rely on labeled datasets to discover patterns of fraudulent conduct, are prevented from being utilized. There are also ethical and legal considerations associated with the labeling of data merely using statistical means. This is because fraud is considered a criminal activity. In addition, fraud is an uncommon occurrence that might have serious repercussions, which renders conventional performance measurements such as accuracy and precision unsuitable. To evaluate fraud detection algorithms, it is more reasonable to use measurements such as recall and specificity.

Through an examination of the phrasing and emotional content of crowdfunding campaigns, language processing models have demonstrated that they have the potential to detect fraudulent activity. The findings of research conducted by Perez et al. (2022) and Cumming et al. (2021) indicated that legitimate campaigns tend to employ language that is more descriptive and informative, whereas fraudulent efforts frequently exhibit language that is more complicated and contains ambiguity. That linguistic analysis has the potential to be used as a method for detecting fraudulent activity in crowdfunding is highlighted here.

Research on fraudulent activity in crowdfunding encompasses a wide range of fields, with a substantial amount of focus being placed on the junction between blockchain technology with applications in the medical field. The use of AI and ML has been implemented to identify abnormalities and outliers in crowdfunding data; nevertheless, the widespread application of these technologies has been hindered by concerns over the availability of data, processing capacity, and privacy.

Through the use of a literature study, the investigation of fraudulent activity in crowdfunding is extremely pertinent for both academics and practitioners. The purpose of a literature review is to offer academics a theoretical framework for understanding how fraud shows itself in various types of crowdfunding, as well as to synthesize the existing body of knowledge and identify research gaps. This makes it possible for future research to be more targeted, to concentrate on areas that have not yet been examined, such as the prevention of fraud in emerging alternative finance models or the application of novel AI/ML technologies in the detection of fraud. In addition to this, it contributes to the development of multidisciplinary approaches, which combine insights from the fields of finance, technology, psychology, and law to produce a comprehensive perspective on fraud in crowdfunding.

The current study can assist practitioners, particularly those involved in platform management or regulation, with actionable insights into fraud detection and prevention. This is especially true for practitioners who have studied the literature. These findings can be utilized by crowdfunding platforms to develop more effective mechanisms for recognizing fraudulent behavior, enhance due diligence procedures, and incorporate more robust data analysis methodologies. To develop or refine policies that provide improved security and transparency in crowdfunding operations, regulators can also benefit from such a study so that they can design or modify policies. Recognizing fraud patterns across industries, particularly in blockchain-based platforms and healthcare, could assist practitioners in mitigating risks and developing more trustworthy systems. Platform managers could also, through partnerships for the production of synthesized datasets, expand the capability to provide deeper insights. The need for expanded research capabilities in crowdfunding, while maintaining platform privacy and competitiveness, presents a significant challenge in the field. While not directly addressing synthesized datasets, Estrin et al. (2021) highlights the importance of both soft and hard information in equity crowdfunding, emphasizing the complex nature of data in this domain. This complexity underscores the potential value of synthesized datasets that could capture nuanced information without compromising sensitive details. Furthermore, the comprehensive review by Joseph & Parasar (2024) on the historical evolution and societal implications of crowdfunding platforms identifies research gaps and establishes a foundation for future investigation. This suggests that there is indeed a need for more robust data sources to address these gaps, which could potentially be fulfilled by synthesized datasets. In the context of innovation and crowdfunding, Bargoni et al. (2022) conducted a bibliometric review that revealed the importance of knowledge flow between different stakeholders in crowdfunding. This finding indirectly supports the argument for shared, synthesized datasets as a means to facilitate this knowledge exchange while protecting individual platform interests. While these sources do not explicitly advocate for the production and sharing of synthesized datasets by crowdfunding platforms, they collectively point to the need for more comprehensive data analysis in crowdfunding research.

The quantity of publications that were examined and the method that was used to obtain the data are both considered to be limitations of this study. Expanding the scope of future study to include various ways of financing in addition to crowdfunding and taking into consideration a wider variety of sources, such as preprints and new databases, may result in the disclosure of more comprehensive information regarding the detection of fraudulent activity. It will be essential to do additional research into the integration of powerful AI and ML models with enhanced data quality and privacy protections to meet the ever-evolving difficulties of fraud in crowdfunding. Additionally, by researching fraud detection procedures in other kinds of alternative finance, such as peer-to-peer lending and initial coin offerings (ICOs), both academics and practitioners can gain a more thorough understanding of the wider landscape of fraud in digital money.

Our study demonstrates that combining human-based literature review with machine learning can effectively address challenging topics, offering valuable insights and inspiration for future research.

References

Akhmadiyev, A., Balgozhina, M., & Tokubayev, K. (2023a). Study of the mechanism for the implementation of crimes related to pyramid schemes and methods of their prevention. *InterEULawEast, 10,* 231-244. URL: https://www.scopus.com/inward/record.uri?eid=2- s2.0-85182649767&doi=10.22598%2fiele.2023.10.2. 12&partnerID=40&md5=2a2bded8542773c18d3b735010eec41b. doi:10.22598/iele.2023.10.2.12.

Akhmadiyev, A., Balgozhina, M., & Tokubayev, K. (2023b). Study of the mechanism for the implementation of crimes related to pyramid schemes and methods of their prevention. *InterEULawEast: Journal for the international and european law, economics and market integrations*, *10*, 231-244.

Alruwaili, A., & Kruger, D. (2020). Crowdfunding with periodic milestone payments using a smart contract to implement fair e-voting. In *Business Information Systems Workshops: BIS 2020 International Workshops, Colorado Springs, CO, USA, June 8-10, 2020, Revised Selected Papers 23* (pp. 61-72). Springer.

Alshater, M. M., Joshipura, M., Khoury, R. E., & Nasrallah, N. (2023). Initial coin offerings: A hybrid empirical review. *Small Business Economics*, *61*, 891-908.

Appio, F. P., Leone, D., Platania, F., & Schiavone, F. (2020). Why are rewards not delivered on time in rewards-based crowdfunding campaigns? an empirical exploration. *Technological Forecasting and Social Change*, *157*, 120069.

Baek, S., Kwon, D., Suh, S. C., Kim, H., Kim, I., & Kim, J. (2021). Clustering-based label estimation for network anomaly detection. *Digital Communications and Networks*, *7*, 37-44.

Bargoni, A., Ferraris, A., Bresciani, S., & Camilleri, M. A. (2022). Crowdfunding and innovation: A bibliometric review and future research agenda. *SSRN Electronic Journal*, . URL: https://api.semanticscholar.org/CorpusID:254522353.

Cardona, L. F., Guzman-Luna, J. A., & Restrepo-Carmona, J. A. (2024). Bibliometric analysis of the machine learning applications in fraud detection on crowdfunding platforms. *Journal of Risk and Financial Management*, *17*, 352.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, *41*, 1-58.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321-357.

Chen, X., & Wei, Z. (2023). Detecting crowdfunding frauds based on textual and imbalanced data. *Data Analysis and Knowledge Discovery*, *7*, 125-135.

Chen, Y., Calabrese, R., & Martin-Barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, *312*, 357-372.

Choi, J., Kim, J., & Lee, H. (2022). Hybrid fraud detection model: Detecting fraudulent information in the healthcare crowdfunding. *KSII Transactions on Internet and Information Systems (TIIS)*, *16*, 1006-1027.

Chou, S.-C., Li, Z.-A., Wang, T., & Yen, J.-C. (2023). How the quality of initial coin offering white papers influences fundraising: Using security token offerings white papers as a benchmark. *Intelligent Systems in Accounting, Finance and Management*, *30*, 3-18. URL: https://onlinelibrary.wiley.com/doi/10.1002/isaf.1527. doi:10.1002/isaf.1527.

Coutrot, I. P., Smith, R., & Cornelsen, L. (2020). Is the rise of crowdfunding for medical expenses in the united kingdom symptomatic of systemic gaps in health and social care? *Journal of Health Services Research & Policy*, *25*, 181-186.

Cumming, D., Hornuf, L., Karami, M., & Schweizer, D. (2021). Disentangling crowdfunding from fraudfunding. *Journal of Business Ethics*, (pp. 1-26).

Cumming, D. J., Johan, S., & Reardon, R. S. (2023). Crowdfunding and intellectual property. *Colo. Tech. LJ*, *22*, 215.

De Caria, R. (2020). Definitions of smart contracts. *The Cambridge handbook of smart contracts, blockchain technology and digital platforms*, (pp. 19-36).

Dheeraj, S., Mahesh, B., Rajput, N. K., Joshi, N. S., & Vasudevan, V. (2022). A large scale medical crowdfunding platform using smart contracts in blockchain. In *2022 IEEE North Karnataka Subsection Flagship International Conference (NKCon)* (pp. 1-6). IEEE.

Drummond, C., Holte, R. C. et al. (2003). C4. 5, class imbalance, and cost sensitivity: why under-sampling beats over-sampling. In *Workshop on Learning from Imbalanced Datasets II, ICML*. Washington, DC volume 11.

Duarte, D. P. (2022). Intermediation risk and conflicts of interest (art 8). In *Regulation on European Crowdfunding Service Providers for Business* (pp. 136-149). Edward Elgar Publishing.

Ellman, M., & Hurkens, S. (2019). Fraud tolerance in optimal crowdfunding. *Economics Letters, 181,* 11-16.

Elmer, G., & Ward-Kimola, S. (2023). Crowdfunding (as) disinformation:'pitching'5g and election fraud campaigns on gofundme. *Media, Culture & Society*, *45*, 578-594.

Estrin, S., Khavul, S., & Wright, M. (2021). Soft and hard information in equity crowdfunding: network effects in the digitalization of entrepreneurial finance. *Small Business Economics*, *58*, 1761 - 1781. URL: https://api.semanticscholar.org/CorpusID: 238846151.

Fang, C., & Stone, W. Z. (2021). An ecosystem for the dairy logistics supply chain with blockchain technology. In *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)* (pp. 1-6). IEEE.

Farajian, M., Lauzon, A. J., & Cui, Q. (2015). Introduction to a crowdfunded public-private partnership model in the united states: policy review on crowdfund investing. *Transportation Research Record*, *2530*, 36-43.

Felix, T. H., & von Eije, H. (2019). Underpricing in the cryptocurrency world: evidence from initial coin offerings. *Managerial Finance*, *45*, 563-578.

Ferretti, R. (2022). Individual portfolio management of loans (art 6). In *Regulation on European Crowdfunding Service Providers for Business* (pp. 113-129). Edward Elgar Publishing.

Gada, S., Dhuri, A., Jain, D., Bansod, S., & Toradmalle, D. (2021). Blockchain-based crowdfunding: A trust building model. In *2021 International conference on artificial intelligence and machine vision (AIMV)* (pp. 1-7). IEEE.

Gaskin, J. E., Keith, M. J., Meservy, T., Twyman, N. W., & Wells, T. (2021). Crowdfunding deception perception: What makes would-be contributors perceive fakeness in crowdfunding campaigns? In *AMCIS*.

Hashemi Joo, M., Nishikawa, Y., & Dandapani, K. (2020). Icos, the next generation of ipos. *Managerial Finance*, *46*, 761-783.

Hernandez Aros, L., Bustamante Molano, L. X., Gutierrez-Portela, F., Moreno Hernandez, J. J., & Rodriguez Barrero, M. S. (2024). Financial fraud detection through the application of machine learning techniques: a literature review. *Humanities and Social Sciences Communications*, *11*, 1-22.

Hooghiemstra, S. N. (2022). The provision of crowdfunding services under the ecspr (art 3). In *Regulation on European Crowdfunding Service Providers for Business* (pp. 68-85). Edward Elgar Publishing.

Huang, L., & Pontell, H. N. (2023). Crime and crisis in china's p2p online lending market: a comparative analysis of fraud. *Crime, Law and Social Change*, *79*, 369-393.

Jadhav, S., Patil, R., Patil, S., Patil, S., & Patil, V. (2023). Ethereum-based decentralized crowdfunding platform. In *International Conference On Innovative Computing And Communication* (pp. 163-175). Springer.

Jeni, L. A., Cohn, J. F., & De La Torre, F. (2013). Facing imbalanced data-recommendations for the use of performance metrics. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 245-251). IEEE.

Jin, D. (2024). Crowdfunding: Is it a double-edged sword on investments? *Business Law Review*, *45*.

Joseph, K., & Parasar, D. (2024). Review of crowdfunding: Historical evolution, societal implications, and architectural perspectives. *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, (pp. 328-334). URL: https: //api.semanticscholar.org/CorpusID:271574644.

Kennedy, R. K., Villanustre, F., Khoshgoftaar, T. M., & Salekshahrezaee, Z. (2024). Synthesizing class labels for highly imbalanced credit card fraud detection data. *Journal of Big Data*, *11*, 38.

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, *9*, 137-163.

Kou, Y., Lu, C.-T., Sirwongwattana, S., & Huang, Y.-P. (2004). Survey of fraud detection techniques. In *IEEE international conference on networking, sensing and control, 2004* (pp. 749-754). IEEE volume 2.

Kumar, A., Lamba, J., Rawal, B. S., Sathiyanarayanan, M., & Alvarez, N. (2023a). Crowdfunding fraud prevention using smart contracts. In *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)* (pp. 185-190). IEEE.

Kumar, K., Vashist, R., & Vashist, P. C. (2023b). A trustful payment system for crowdfunding using blockchain. In *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)* (pp. 774-780). IEEE. URL: https://ieeexplore. ieee.org/document/10085649. doi:10.1109/AISC56616.2023.10085649.

Lacan, C., & Desmet, P. (2017). Does the crowdfunding platform matter? risks of negative attitudes in two-sided markets. *Journal of Consumer Marketing*, *34*, 472-479.

Lau, K. L., Chew, B. C., Maliki, A., & Hafizuddin, H. (2018). Crowdfunding for academic projects: A case in public universities of malaysia. *Journal of Advanced Manufacturing Technology (JAMT), 12,* 273-284.

Lazaroiu, G., Bogdan, M., Geamanu, M., Hurloiu, L., Luminita, L., & Stefanescu, R. (2023). Artificial intelligence algorithms and cloud computing technologies in blockchain-based fintech management. *Oeconomia Copernicana*, *14*, 707-730.

Lee, S., Shafqat, W., & Kim, H.-c. (2022). Backers beware: Characteristics and detection of fraudulent crowdfunding campaigns. *Sensors*, *22*, 7677.

Li, G., Wang, S., & Feng, Y. (2024). Making differences work: Financial fraud detection based on multi-subject perceptions. *Emerging Markets Review*, *60*, 101134.

Li, X., Sun, X., Meng, Y., Liang, J., Wu, F., & Li, J. (2019). Dice loss for data-imbalanced NLP tasks. *arXiv preprint arXiv:1911.02855*, .

Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal loss for dense object detection. In *proceedings of the IEEE Conference on Computer Vision and Pattern recognition* (pp. 2980-2988).

Liu, B., & Tsoumakas, G. (2020). Dealing with class imbalance in classifier chains via random undersampling. *Knowledge-Based Systems*, *192*, 105292.

Liu, R., Huang, J., & Zhang, Z. (2023). Tracking disclosure change trajectories for financial fraud detection. *Production and Operations Management*, *32*, 584-602.

Lu, S., Xu, X., Wang, H., Zhao, J., & Wu, Z. (2018). Detecting systemically important platforms in p2p market of china. In *2018 15th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 1-7). IEEE.

Maqbool, O., & Babri, H. A. (2006). Automated software clustering: An insight using cluster labels. *Journal of Systems and Software*, *79*, 1632-

1648.

Mayer, L. H. (2022). Regulating charitable crowdfunding. *Ind. LJ*, *97*, 1375.

Mbarek, S., & Trabelsi, D. (2020). Crowdfunding without crowd-fooling: prevention is better than cure. In *Corporate Fraud Exposed: A Comprehensive and Holistic Approach* (pp. 221-238). Emerald Publishing Limited.

Midha, M., Bhardwaz, S., Godha, R., Mehta, A. R., Parida, S. K., & Panda, S. K. (2023). Blockchain-powered crowdfunding: Assessing the viability, benefits, and risks of a decentralized approach. In *International Conference on Data & Information Sciences* (pp. 179-189). Springer.

Moreo, A., Esuli, A., & Sebastiani, F. (2016). Distributional random oversampling for imbalanced text classification. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 805-808).

Moslehi, F., Haeri, A., & Gholamian, M. R. (2020). A novel selective clustering framework for appropriate labeling of clusters based on k-means algorithm. *Scientia Iranica*, *27*, 2621-2634.

Moysidou, K., & Hausberg, J. P. (2020). In crowdfunding we trust: A trust-building model in lending crowdfunding. *Journal of Small Business Management*, *58*, 511-543.

Naik, P. G., & Oza, K. S. (2023). Leveraging the power of blockchain technology for building a resilient crowdfunding solution. *Procedia Computer Science*, *230*, 11-20.

Olvedi, T. (2022). The liquidity aspects of peer-to-peer lending. *Studies in Economics andFinance*, *39*, 45-62.

van Otterloo, S. (2022). Measuring project success: the fulfillment rate of crowdfunded projects on kickstarter. *Computers and Society Research Journal,(4)*, .

Pandey, S., Goel, S., Bansla, S., & Pandey, D. (2019). Crowdfunding fraud prevention using blockchain. In *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1028-1034). IEEE.

Parmar, D., Kori, P., Chauhan, O. S., & Wadmare, J. (2022). A review: A blockchain based crowdfunding decentralized application. In *2022 5th International Conference on Advances in Science and Technology (ICAST)* (pp. 435-439). IEEE.

Perez, B., Machado, S., Andrews, J., & Kourtellis, N. (2022). I call bs: Fraud detection in crowdfunding campaigns. In *Proceedings of the 14th ACM Web Science Conference 2022* (pp. 1-11).

Petrov, L. F., & Emelyanova, E. S. (2021). The crowdfunding: Financial flows and risks. In *CEUR Workshop Proceedings* (pp. 41-51). volume 2830.

Pierce-Wright, C. H. (2016). State equity crowdfunding and investor protection. *Wash. L. Rev.*, *91*, 847.

Pinjarkar, V. U., Pinjarkar, U. S., Bhor, H. N., & Rathod, S. (2023). Crowdfunding campaigns web application using metamask. In *2023 6th International Conference on Advances in Science and Technology (ICAST)* (pp. 217-222). IEEE.

Prashar, A., & Gupta, P. (2024). How to build trust in gen y in online donation crowdfunding: an experimental study. *Behaviour & Information Technology*, *43*, 677-694.

Rajarajeswari, S., Karthik, K., Divyasri, K., Anvith, & Singhal, R. (2023). Open permissioned blockchain solution for private equity funding using a global, cross-cloud network blockchain platform. In *International Conference on Power Engineering and Intelligent Systems (PEIS)* (pp. 289-297). Springer.

Rauber, A. (1999). Labelsom: On the labeling of self-organizing maps. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No. 99CH36339)* (pp. 3527-3532). IEEE volume 5.

Renwick, M. J., & Mossialos, E. (2017). Crowdfunding our health: economic risks and benefits. *Social Science & Medicine*, *191*, 48-56.

Robock, Z. (2014). The risk of money laundering through crowdfunding: A funding portal's guide to compliance and crime fighting. *Mich. Bus. & Entrepreneurial L. Rev.*, *4*, 113.

Rodriguez-Garnica, G., Gutierrez-Urtiaga, M., & Tribo, J. A. (2024). Signaling and herding in reward-based crowdfunding. *Small Business Economics*, (pp. 1-28).

Rosli, F. I., & Shahida, S. (2019). Adressing the principal-agent problem in equity crowdfunding in malaysia. *International Journal of Islamic Economics and Finance Research*, *2*, 26-40.

Saadat, M. N., Halim, S. A., Osman, H., Nassr, R. M., & Zuhairi, M. F. (2019). Blockchain based crowdfunding systems. *Indonesian Journal of Electrical Engineering and Computer Science*, *15*, 409-413.

Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications*, *40*, 5916-5923.

Sahu, M., Gangaramani, A., & Bharambe, A. (2021). Secured crowdfunding platform using blockchain. In *Intelligent Computing and Networking: Proceedings of IC-ICN 2020* (pp. 27-39). Springer.

Salido-Andres, N., Rey-Garcia, M., Alvarez-Gonzalez, L. I., & Vazquez-Casielles, R. (2022). When the winner takes it all: online campaign factors influencing the success of donation-based crowdfunding for charitable causes. *International Review on Public and Nonprofit Marketing*, *19*, 763-780.

Sarmah, M., Saxena, S., & Mukherjee, S. (2022). A decentralized crowdfunding solution on top of the ethereum blockchain. In *2022 IEEE Silchar Subsection Conference (SILCON)* (pp. 1-6). IEEE.

Schwartz, A. A. (2012). Crowdfunding securities. *Notre Dame L. Rev.*, *88*, 1457.

Schwartz, A. A. (2018). The gatekeepers of crowdfunding. *Wash. & Lee L. Rev.*, *75*, 885.

Sewaid, A., Garcia-Cestona, M., & Silaghi, F. (2021). Resolving information asymmetries in financing new product development: The case of reward-based crowdfunding. *Research Policy*, *50*, 104345.

Shafqat, W., & Byun, Y.-C. (2019). Topic predictions and optimized recommendation mechanism based on integrated topic modeling and deep neural networks in crowdfunding platforms. *Applied Sciences*, *9*, 5496.

Shneor, R., Wenzlaff, K., Boyko, K., Baah-Peprah, P., Odorovic, A., & Okhrimenko, O. (2023). The european crowdfunding market report 2023. *Shneor, R., Wenzlaff, K., Boyko, K., Baah-peprah, P., Odorovic, A., and Okrimenko, O.(2024). The European Crowdfunding Market Report*, .

Siering, M., Koch, J.-A., & Deokar, A. V. (2016a). Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. *Journal of Management Information Systems*, *33*, 421-455. URL: https: //www.tandfonline.com/doi/full/10.1080/07421222.2016.1205930. doi:10.1080/07421222.2016.1205930.

Siering, M., Koch, J.-A., & Deokar, A. V. (2016b). Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts. *Journal of Management Information Systems*, *33*, 421-455.

Siswoyo, K. R., Miftah, A., Karnadi, R., & Halim, E. (2023). The role of social presence, social media attribute and empathetic concern to donation intention in indonesia donation-based crowdfunding. In *Proceeding - International Conference on Information Technology and Computing 2023, ICITCOM 2023* (pp. 113-118). Institute of Electrical and Electronics Engineers Inc. URL: https://www.scopus.com/inward/record.uri?eid=2- s2.0- 85187223858&doi= 10.1109%2fICITCOM60176.2023.10442089&partnerID=40&md5=817f542006007bd6453444f95b07fef4. doi:10.1109/ ICITCOM60176.2023.10442089.

Sloan, C. E., Campagna, A., Tu, K., Doerstling, S., Davis, J. K., & Ubel, P. A. (2023). Online crowdfunding campaigns for diabetes- related expenses. *Annals of Internal Medicine, 176,* 1012-1014.

Stack, P., Feller, J., O'Reilly, P., Gleasure, R., Li, S., & Cristoforo, J. (2017). Managing risk in business centric crowdfunding platforms. In *Proceedings of the 13th International Symposium on Open Collaboration* (pp. 1-4).

Sun, H., Li, J., & Zhu, X. (2023). Financial fraud detection based on the part-of-speech features of textual risk disclosures in financial reports. *Procedia Computer Science*, *221*, 57-64.

Sureshbhai, P. N., Bhattacharya, P., & Tanwar, S. (2020). Karuna: A blockchain-based sentiment analysis framework for fraud cryptocurrency schemes. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)* (pp. 1-6). IEEE.

Teichmann, F., Boticiu, S. R., & Sergi, B. S. (2024). Compliance risks for crowdfunding. a neglected aspect of money laundering, terrorist financing and fraud. *Journal of Financial Crime*, *31*, 575-582.

Teichmann, F. M. J., Boticiu, S. R., & Sergi, B. S. (2023). Compliance concerns in sustainable finance: an analysis of peer-to-peer (p2p) lending platforms and sustainability. *Journal of Financial Crime*, .

Van Osnabrugge, M. (2000). A comparison of business angel and venture capitalist investment procedures: an agency theory-based analysis. *Venture Capital: An international journal of entrepreneurial finance*, *2*, 91-109.

Wang, H. (2019). Detection of fraudulent users in p2p financial market. *arXiv preprint arXiv:1910.02010*, .

Wang, X., & Wang, L. (2019). The role of charitable crowdfunding platforms on poverty alleviation. *Journal of Advances in Information Technology*, *10*, 72-76. URL: https://www.scopus.com/inward/record.uri?eid=2- s2.0-85147179566& doi=10.12720%2fjait.10.2.72-76&partnerID=40&md5=79b3902a405799e3a688307bbef2f67d. doi:10.12720/jait. 10.2.72-76.

Wenzlaff, K. (2023). Crowdfunding for science and teaching in higher education: Status quo and research agenda. *Crowdfunding in Higher Education Institutions: Theory and Best Practices*, (pp. 17-29).

Wenzlaff, K., Odorovic, A., Kleverlaan, R., & Ziegler, T. (2021). Assessing the maturity of crowdfunding and alternative finance markets. *Crowdfunding in the Public Sector: Theory and Best Practices*, (pp. 65-75).

Wenzlaff, K., Odorovic, A., Riethmuller, T., & Wambold, P. (2022). On the merits of the key investment information sheet in the ecspr (arts 23-24 and annex i). In *Regulation on European crowdfunding service providers for business* (pp. 310-349). Edward Elgar Publishing.

Wessel, M., Thies, F., & Benlian, A. (2022). The role of prototype fidelity in technology crowdfunding. *Journal of Business Venturing*, *37*, 106220.

Wu, X., Dinger, H., & Yuksel, S. (2022). Analysis of crowdfunding platforms for microgrid project investors via a q-rung orthopair fuzzy hybrid decision-making approach. *Financial Innovation*, *8*, 52.

Wu, X., Nguyen, T. T., Zhang, D. C., Wang, W. Y., & Luu, A. T. (2023a). Fastopic: Pretrained transformer is a fast, adaptive, stable, and transferable topic model. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Wu, X., Pan, F., & Luu, A. T. (2023b). Towards the topmost: A topic modeling system toolkit. *arXiv preprint arXiv:2309.06908*, .

Xu, J. J., Chen, D., Chau, M., Li, L., & Zheng, H. (2022). Peer-to-peer loan fraud detection: Constructing features from transaction data. *MIS quarterly*, *46*.

Xu, J. J., Lu, Y., & Chau, M. (2015). P2p lending fraud detection: A big data approach. In *Intelligence and Security Informatics: Pacific Asia Workshop, PAISI 2015, Ho Chi Minh City, Vietnam, May 19, 2015. Proceedings* (pp. 71-81). Springer.

Xu, Y., Li, Q., Zhang, C., Tan, Y., Zhang, P., Wang, G., & Zhang, Y. (2023). A decentralized trust management mechanism for crowdfunding. *Information Sciences*, *638*, 118969.

Yeon, A. L., Yaacob, N., Hussain, M. A., & Md Ismail, C. T. (2022). Equity crowdfunding vs cybercrime: A legal protection. *BiLD Law Journal*, *7*, 139-150.

Yotsawat, W., Wattuya, P., & Srivihok, A. (2021). A novel method for credit scoring based on cost-sensitive neural network ensemble. *IEEE Access*, *9*, 78521-78537.

Zenone, M., & Snyder, J. (2019). Fraud in medical crowdfunding: A typology of publicized cases and policy recommendations. *Policy & Internet*, *11*, 215-234.

Ziegler, T., Shneor, R., Wenzlaff, K., Wang, B., Kim, J., Paes, F. F. d. C., Suresh, K., Zhang, B. Z., Mammadova, L., & Adams, N. (2021). The global alternative finance market benchmarking report. *Available at SSRN 3771509*, .

Zilgalvis, P. (2014). The need for an innovation principle in regulatory impact assessment: the case of finance and innovation in europe. *Policy & Internet, 6,* 377-392.

Zkik, K., Sebbar, A., Fadi, O., Kamble, S., & Belhadi, A. (2024). Securing blockchain-based crowdfunding platforms: an integrated graph neural networks and machine learning approach. *Electronic Commerce Research*, *24*, 497-533.

Appendix

Table A1: Table reporting all the articles that have been examined to conduct the research and highlighting their main features

Study Authors	Fraud Definition?	Datasets Sources	Label Fraud? If so, how?	Fraud Detection Methods	ML Methods Used in the Study	Validation Methods Used	Main Contributions	Main Limitations
Sarmah et al. (2022)	No fraud definition	No dataset, proposes new architecture	No labeling	Blockchain transparency for accountability	No ML methods	No validation	Proposes decentralized platform on Ethereum with smart contracts	Conceptual, no deployment or testing
Xu et al. (2023)	Behavior-based fraud	No dataset, simulations	Yes, linked to auditor behavior	Thresholds, penalties, independent audits	No ML methods	Simulations in Truffle framework	Blockchain trust management mechanism	Scalability, audit integrity concerns
Parmar et al. (2022)	No fraud definition	No dataset, conceptual blockchain proposal	No labeling	DAOs for transparency in fund usage	No ML methods	No validation	Highlights transparency via DAOs in crowdfunding	Lacks empirical testing
Dheeraj et al. (2022)	Defines fraud as significant threat to crowdfunding	No dataset, proposes new platform	No labeling	Smart contracts, voting system to prevent fraud	No ML methods	No validation	Secure medical crowdfunding via blockchain	Legal and technological challenges
Naik & Oza (2023)	Yes, blockchain-based fraud prevention	Theoretical framework	No specific labeling	Smart contracts and blockchain transparency	No ML methods	Conceptual validation	Resilient crowdfunding architecture using blockchain	Limited practical implementation details
Kumar et al. (2023a)	Yes, smart contract based fraud detection	Experimental setup	Smart contract validation	Automated fraud prevention through contracts	No ML methods	Conference prototype	Smart contract based fraud prevention system	Limited scalability testing
Fang & Stone (2021)	No fraud definition	Conceptual paper, no dataset	No labeling	Blockchain to ensure transparency in supply chain	No ML methods	No validation	Blockchain proposal for dairy logistics transparency	Complex IoT integration, adoption challenges
Wu et al. (2022)	No fraud definition	No dataset, fuzzy sets used	No labeling	Fuzzy sets to assess platform reliability	No ML methods	Sensitivity analysis	Fuzzy model for microgrid crowdfunding platforms	Limited to microgrid sector
Lazaroiu et al. (2023)	No fraud definition	Literature review	Fraud as anomaly in transactions	AI algorithms for anomaly detection	No ML methods specified	No validation	Integration of AI in fin- tech fraud detection	Focus on recent studies only
Lee et al. (2022)	Fraud as deceptive crowdfunding campaigns	Kickstarter dataset	Yes, manual and automated labeling	Text and behavioral analysis of campaigns	Supervised learning	Cross-validation, accuracy metrics	Combines text and behavior analysis for fraud detection	Kickstarter-specific dataset, manual bias

Alruwaili & Kruger (2020)	Fraud in evoting systems for crowdfunding	No dataset	No labeling	Blockchain-based evoting for milestone payments	No ML methods	No validation	Blockchain proposal for secure milestone payments	No real-world application or testing
Hashemi Joo et al. (2020)	Fraud linked to ICO transparency	Literature review	No labeling	Blockchain transparency	No ML methods	No validation	Reviews risks and opportunities of ICO fraud	Conceptual, no empirical data
Farajian et al. (2015)	No fraud definition	No dataset, conceptual	No labeling	Public-private crowdfunding model	No ML methods	No validation	Proposes public-private partnership in crowdfunding	Theoretical framework, no real-world testing
Coutrot et al. (2020)	No fraud definition	Conceptual paper, UK healthcare crowdfunding	No labeling	Analysis of crowdfunding in UK healthcare	No ML methods	No validation	Gaps in healthcare crowdfunding in UK	UK-specific, lacks broader scope
Saadat et al. (2019)	Fraud in unregulated crowdfunding	No dataset provided	Fraud inferred by project failure	Blockchain transparency, smart contracts	No ML methods	Unit and integration testing on Mocha	Blockchain-based system for Malaysian perspective	No empirical data, theoretical model
Prashar & Gupta (2024)	Fraud as financial data theft	Data from 9 countries	No fraud labeling	Blockchain integration with secure ledgers	No ML methods	No validation methods mentioned	Secure crowdfunding via smart contracts	Limited to conceptual model
Gada et al. (2021)	No fraud definition	No dataset provided	No fraud labeling	Blockchain-based trust building	No ML methods	No validation methods mentioned	Proposes trust model for crowdfunding using blockchain	Lacks empirical validation
Midha et al. (2023)	No fraud definition	No dataset provided	No fraud labeling	Blockchain and decentralized platforms	No ML methods	No validation methods mentioned	Proposes decentralized crowdfunding approach	Conceptual, no real-world application
Teichmann et al. (2024)	Fraud in crowdfunding as a risk for money laundering	No dataset provided	No fraud labeling	Compliance framework for risk prevention	No ML methods	No validation methods mentioned	Regulatory framework for fraud risk in crowdfunding	Focuses on compliance, lacks empirical data
Renwick & Mossialos (2017)	Fraud in medical crowdfunding as misuse of funds	No dataset provided	No fraud labeling	Trust and transparency in medical crowdfunding	No ML methods	No validation methods mentioned	Proposes transparent models for medical crowdfunding	Theoretical, lacks empirical testing
Schwartz (2012)	Fraud in medical crowdfunding as a form of scam	No dataset provided	No fraud labeling	Trust in crowdfunding for medical causes	No ML methods	No validation methods mentioned	Examines fraud risks in medical crowdfunding	Conceptual, no real-world application
Elmer & Ward-Kimola (2023)	Disinformation in crowdfunding	Six election fraud and 5G campaigns	No fraud labeling	Language analysis for disinformation detection	No ML methods	No validation methods mentioned	Analyzes crowdfunding disinformation campaigns	Focuses on disinformation rather than fraud

Pinjarkar et al. (2023)	No fraud definition	No dataset provided	No fraud labeling	Crowdfunding web app using blockchain	No ML methods	No validation methods mentioned	Proposes blockchainbased crowdfunding app	No empirical data, theoretical model
Gaskin et al. (2021)	Fraud as deception in crowdfunding campaigns	818 evaluations of COVID-19 campaigns	Fraud labeling via NLP measures	NLP and behavioral analysis of crowdfunding campaigns	No ML methods	No validation methods mentioned	Identifies deception patterns in crowdfunding campaigns	Focuses on signals, lacks empirical validation
Pandey et al. (2019)	Fraud in blockchain crowdfunding systems	Rinkeby test network	No fraud labeling	Blockchain transparency and smart contracts	No ML methods	No validation methods mentioned	Proposes blockchain-based fraud prevention in crowdfunding	Limited to Ethereum network
Zenone & Snyder (2019)	Fraud in medical crowdfunding as impersonation	News articles from LexisNexis and GoFraudMe	Fraud labeled via thematic analysis	Case analysis of medical crowdfunding fraud	No ML methods	Independent review, thematic analysis	Provides typology of medical crowdfunding fraud	Limited to specific platforms
Ellman & Hurkens (2019)	Fraud tolerance in optimal crowdfunding	No dataset, theoretical study	No fraud labeling	Mathematical modeling of fraud tolerance	No ML methods	Theoretical validation	Demonstrates optimal fraud tolerance in crowdfunding	Theoretical, no empirical data
Choi et al. (2022)	Fraud in healthcare crowdfunding as misrepresentation	GoFundMe dataset of 10,012 campaigns	Labeled via expert and textual analysis	Hybrid fraud detection using LDA and CF	LDA, CF	Comparative analysis and ML evaluation	Develops hybrid model for fraud detection in healthcare crowdfunding	Limited to GoFundMe, lacks generalizability
Perez et al. (2022)	Fraud as misrepresentation in crowdfunding	Crowdfunding platforms like GoFundMe	Fraud labeled via manual annotation	Feature extraction and supervised classification	Ensemble classifier	Random split validation	Proposes ML model for detecting fraud in crowdfunding campaigns	Manual annotation biases, dataset limitations
Alshater et al. (2023)	Fraud in ICOs as misrepresentation	ICO white papers from 2017-2020	No fraud labeling	Text analysis of ICO white papers	No ML methods	Logistic regression	Analyzes fraud risks in ICO white papers	Limited dataset, lacks empirical data
Sureshbhai et al. (2020)	Fraud in cryptocurrency as Ponzi schemes	Elliptic dataset	Fraud labeled using Bitcoin transaction data	Sentiment analysis and LSTM	LSTM	Train-test split for validation	Proposes LSTM model for detecting Ponzi schemes	Limited to cryptocurrency fraud
Stack et al. (2017)	Fraud in business-centric crowdfunding as misrepresentation	No dataset provided	Fraud labeled via platform policies	Self-governance, regulatory compliance	No ML methods	No validation methods mentioned	Proposes self-governance for crowdfunding platforms	Theoretical, lacks empirical data
Rajarajeswari et al. (2023)	Fraud in private equity funding as misuse of funds	No dataset provided	Fraud labeled via discrepancies in cap tables	Blockchain transparency for secure cap tables	No ML methods	Blockchain validation mechanisms	Proposes blockchain for private equity fraud prevention	Scalability and regulatory concerns

Xu et al. (2015)	Fraud in P2P lending as misrepresentation	Chinese P2P lending platforms	Fraud labeled via transaction patterns	Data mining and anomaly detection	Decision Trees, SVM, Neural Networks	Cross-validation	Combines ML methods for P2P lending fraud detection	Data quality and scalability issues
Mayer (2022)	Fraud in charitable crowdfunding as misrepresentation	No dataset provided	Fraud labeled via campaign discrepancies	Regulatory recommendations for fraud prevention	No ML methods	Comparative analysis of regulations	Proposes regulatory measures for charitable crowdfunding fraud	Theoretical, lacks empirical data
Sahu et al. (2021)	Fraud in crowdfunding as misuse of funds	Crowdfunding platforms data	Fraud labeled via discrepancies in campaign updates	Blockchain and smart contracts for fraud prevention	Decision Trees, Random Forest, SVM, Neural Networks	Cross-validation	Proposes blockchainbased fraud detection using smart contracts	Scalability issues, theoretical model
Zkik et al. (2024)	Fraud in blockchain based crowd funding as cyber-attacks	Crowdfunding platforms data	Fraud labeled via transaction anomalies	Graph Neural Networks (GNN) and ML models for anomaly detection	GNNs, Random Forest, SVM, Neural Networks	Cross-validation, performance metrics	Proposes GNN and ML integration for crowdfunding fraud detection	Computational complexity, data dependency
Rodriguez- Garnica et al. (2024)	No explicit fraud focus	Kickstarter data	No fraud labeling	Analyzes signaling and herding behaviors in crowdfunding	No ML methods	Empirical analysis of Kickstarter data	Provides insights into herding behavior in crowdfunding	Limited fraud focus, platform-specific
Pierce-Wright (2016)	Fraud in equity crowdfunding as misrepresentation	Historical data on state crowdfunding exemptions	No fraud labeling	Regulatory oversight, issuer requirements	No ML methods	Historical and regulatory analysis	Proposes regulatory measures to protect investors in equity crowdfunding	Theoretical, lacks empirical data
Appio et al. (2020)	No explicit fraud definition	Kickstarter data	No fraud labeling	Text mining to detect delays and possible fraud	Text mining	None provided	Examines delays in reward-based crowdfunding projects	Limited focus on fraud detection
Petrov & Emelyanova (2021)	No fraud definition provided	No dataset provided	No fraud labeling	Financial risk analysis in crowdfunding	No ML methods	No validation methods mentioned	Analyzes financial flows and risks in crowdfunding	Lacks empirical data and fraud focus
Shafqat & Byun (2019)	Fraudulent campaigns via threatening language	Crowdfunding platforms data	Fraud labeled via language analysis	Text mining and DNN for fraud detection	DNN	Performance evaluation metrics	Proposes text-based fraud detection in crowdfunding	Limited dataset, no empirical validation
	Fraud in charitable crowdfunding as multiple fake projects	Leijuan platform data	Fraud labeled via agent behavior	Pattern analysis of project creation	No ML methods	No validation methods mentioned	Identifies fraud patterns in donation-based crowdfunding	Limited to specific platform

Siswoyo et al. (2023)	Fraud in donation-based crowdfunding as a risk	Survey data from 144 respondents	No fraud labeling	Social presence and empathetic concern	No ML methods	SmartPLS for validation	Examines social factors affecting donation behavior	No fraud focus, limited dataset
Akhmadiyev et al. (2023b)	Fraud in pyramid schemes	Case studies on pyramid schemes	Fraud labeled via characteristics of pyramid schemes	Legal and theoretical analysis of fraud prevention	No ML methods	Comparative legal analysis	Proposes international regulations for pyramid schemes	Theoretical, lacks empirical data
Zilgalvis (2014)	No fraud focus	No dataset provided	No fraud labeling	Regulatory impact assessment for innovation	No ML methods	No validation methods mentioned	Proposes innovation principle in regulatory impact	No fraud focus, conceptual
Chou et al. (2023)	Fraud in ICOs as misrepresentation	ICO white papers from 2017-2020	No explicit fraud labeling	Textual analysis of ICO and STO white papers	No ML methods	Logistic regression for validation	Analyzes quality of ICO white papers to prevent fraud	Limited dataset, lacks empirical validation