IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 17 February 2025, accepted 25 February 2025, date of publication 3 March 2025, date of current version 10 March 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3547396

RESEARCH ARTICLE

Fraud Detection on Crowdfunding Platforms Using Multiple Feature Selection Methods

SEUNGHUN LEE, HEEMIN PARK, AND HYUN-CHUL KIM, (Member, IEEE)
Department of Software, Sangmyung University, Cheonan 31066, South Korea

Corresponding author: Hyun-Chul Kim (hyunchulk@gmail.com)

This research was funded by a 2022 Research Grant from Sangmyung University.

ABSTRACT In recent years, crowdfunding has emerged as an alternative funding source for startups and emerging businesses, experiencing significant growth. However, this growth has also led to an increase in fraudulent activities. Despite the potential for fraud in the realm of crowdfunding, there is limited knowledge of the phenomenon due to a lack of data on actual instances of fraudulent campaigns. In this paper, we aim to address this deficiency by collecting and analyzing publicly accessible web and social media data from a hundred fraudulent crowdfunding projects. In order to identify and comprehend the distinguishing characteristics of fraudulent campaigns, we first propose 1) using a wide variety of characteristics of campaign projects and project creators, including their profiles, behavior, social traits, and language; then, 2) we propose to use and combine three well-known multiple feature selection methods, which are based on Correlation-based Feature Selection (CFS), Pearson Correlation Coefficient (PCC), and Information Gain (IG), to identify representative features of fraudulent campaigns. Our approach identifies 10 commonly selected key features of fraudulent crowdfunding campaigns, three of which are new, original findings. We provide and discuss our findings and interpretations on the 10 commonly selected key features in relation to previous studies, based on which we construct a fraud detection model with 82.04% accuracy. We also employ Shapley Additive ExPlanations (SHAP) to interpret the fraud detection model, explaining the importance of each feature.

INDEX TERMS Fraud detection, crowdfunding, machine learning, natural language processing, feature selection, correlation-based feature selection, Pearson correlation coefficient, information gain, shapley additive explanations, explainability.

I. INTRODUCTION

In recent years, the emergence of crowdfunding has revolutionized the fundraising landscape, presenting itself as a potent alternative to traditional methods of acquiring investments. Crowdfunding harnesses the power of the masses, by appealing to a vast network of people through advertising project ideas online. This novel method has become an invaluable source of support for individuals, small businesses, startups, and industries that are grappling with the financial hurdles of starting up. Over the past decade, crowdfunding has proven to be a reliable and sustainable means of securing investment, standing its ground against the traditional venture capital model. The worldwide crowdfunding market is growing rapidly each year, with $5.9 billion raised in 2021 and $6.5 billion raised in 2022 [1].

Despite its popularity and success, crowdfunding faces significant obstacles, including the high risk of fraud [2]. The structure of crowdfunding, such as simplicity in exemplifying the idea, ease of use, flexibility of requirements, lack of legal resources for investors, and absence of financial intermediaries, can cause increased fraudulent activity [3], [4]. In conventional forms of entrepreneurial financing, investors are financial institutions with access to comprehensive records of the founder's (i.e., project creator's) credit history, academic or professional background, and the assistance of experts in the corresponding or closely related

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Liu.

field who can evaluate the quality of expected outcomes, the team, and the probability of success [4], [5]. However, a setting like the current platforms for crowdfunding, in which individual amateurs choose which projects to fund based on the limited information about project campaigns and founders, is susceptible to fraud and abuse. Currently, crowdfunding platforms raise funds without legal sanctions against creators who promise to deliver promised rewards on time. Consequently, it is always possible that fraudsters may abuse the whole system and investors' trust, which makes them vulnerable to fraud. The "Kobe red beef jerky" campaign by Magnus Inc., which claimed to offer fresh Kobe beef-based jerky from Japan and produced bogus user experiences proving that they enjoy the taste, is a well-known attempt of fraudulent crowdfunding [3]. It almost achieved $120,309, roughly 50 times the campaign's original funding goal, from 3,252 backers (i.e., investors) in less than a month. Fortunately, Kickstarter spotted this scam at the last minute of fundraising time, after a documentary film campaign named "kickstarted" raised questions and doubts about the legitimacy of the campaign in a Reddit post [6].

Crowdfunding users typically use sites such as reddit.com [7], kickscammed.com [8], and the Facebook group "crowdfunding projects that never delivered" [9] (where hundreds of suspicious campaigns are reported and discussed by victims) to exchange information, discuss projects, and report fraud. This is a strong indication that victims are experiencing worry and disruption and a sign that the general public must be protected from the impending invasion of deceivers. As an example, this recent increase in alleged fraud has also resulted in federal and state-level legal actions. The Federal Trade Commission (FTC) has initiated the first legal enforcement action against a crowdfunding campaign titled *"The Doom That Came To Atlantic City!"*, showing the willingness of the FTC to protect consumers utilizing this emerging and developing financial technology [10].

The rationale behind regulations and legal actions is straightforward: fraud must be curbed to allow crowdfunding models to flourish as an alternative, viable, and long-lasting source of funding for start-up businesses, which implies that unwary contributors, donors, and investors must be protected [11]. Although crowdfunding frauds are a new and serious threat to investors, their research has been hampered by a lack of measurement data collected from a good number of actual fraud cases. Our research fills this gap (i) by collecting and analyzing a hundred fraudulent crowdfunding campaigns, (ii) using a wide range of characteristics of the campaigns and their creators, and then (iii) combining multiple feature selection methods. We highlight the key contributions of our study:

a. We collected and analyzed hundreds of fraudulent campaigns from one of the most popular and publicly available crowdfunding sites, Kickstarter.com. We begin with 90 features[1] inspired by previous related studies [4], [12], [13], [14], which can be categorized as follows: (i) campaign information like the number of videos and funding goal, (ii) campaign creators' profile and behavior information like the number of backed or created projects and the number of comments they have left on the other campaigns, and (iii) linguistic features that have been proved useful in deception and fraud detection, like the use of pronouns, cognitive process words (i.e., "think", "because", "know"), and Readability [15], [16], [17], [18], extracted from all *Campaign*, *Updates*, and *Comments* sections available on the platform for fundraisers to pitch and interact with backers.

b. We strategically propose to use and combine three well-known multiple feature selection methods, which are Correlation-Based Feature Selection [19], Pearson correlation coefficient [20], and Information Gain [19], in order to identify representative characteristics of fraudulent campaigns. In particular, we use the intersection strategy of the three feature selection methods, to filter out less or unrepresentative features [21]. As a result, we successfully identified 10 commonly selected key features of fraudulent campaigns, three of which are new, original findings. We found that 3 (out of those 10) features from the *Comments* section are the most useful for detecting frauds, followed by those from the Campaign & Creator's information (3 out of 10), and then those from the *Updates* section (4 out of 10). We build our fraud detection model with the 10 key features, achieving 82.04% classification accuracy.

c. We discuss our findings and their interpretations on the key characteristics of fraudulent projects and their founders, in relation to previous studies. In particular, the following 3 characteristics out of those 10 are our own original findings: (i) the number of videos in the *Updates* section, which seems to be daunting for fraudsters to produce and add their owns, turned out to be a good predictor for fraud detection. We also found that fraudsters tend to use words of (ii) insight (i.e., "think", "know") and (iii) causation (i.e., "because", "hence") more often in *Comments* section, attempting to (a) make false excuses to justify the lack of progress or unpreparedness, and delay of reward delivery schedules, as well as (b) present subjective beliefs or claims rather than objective, verifiable facts. We interpret these results in more detail in Section IV. The other key features of fraudulent campaigns and their creators we found corroborating previous studies include: the smaller number of backed campaigns, lower typo ratio, less use of location words, first-person plural nouns, external WWW links, and email contacts.

d. We employ SHAP to interpret the fraud detection model by quantifying the contribution of each feature to the detection process. Through this analysis, we validate that the features from the comments section play a crucial role in

[1] In this paper, a total of 90 features are utilized, encompassing 18 features from the project and creator information in Table 3, as well as 24 linguistic characteristics each from the *Campaign*, *Updates*, and *comments* sections, which can be observed in Tables 6, and 7, respectively.

detecting fraudulent projects, with two of our original finding features (i.e., insight and causal words) standing out as key contributors, ranking first and fourth in importance.

This paper is organized as follows: In Section II, reviewing related work. Our dataset and methodology are detailed in Section III. The results of extracted key features and our model for detecting fraudulent campaigns, followed by SHAP-based interpretability analysis, are discussed and presented in Section IV. Finally, the paper concludes in Section V.

II. RELATED WORK

A. CROWDFUNDING

The majority of research on crowdfunding has focused on identifying the variables that influence fundraising success or failure and predicting success [22], [23], [24], [25], [26], [27], [28], [29], [30]. Mollick explored the relationship between project features and their impact on project success [28]. He found that certain static features of projects, such as the presence of a video or spelling errors and grammar, and social network factors, such as the number of Facebook friends of the project creator, are highly correlated with the success of a crowdfunding campaign [28]. Furthermore, Mollick uncovered that a considerable 9% of funded projects failed to deliver their rewards [31]. Greenberg et al. demonstrated to be able to predict the success or failure of a project with 68% accuracy using Support Vector Machines (SVM) at the time of launch [26]. It has been known that both online social networks (i.e., related tweets or retweets to a campaign) [22] and offline social relationships such as friends or family [32] influence the success of online campaigns. Features associated with project content, such as having a high-quality video [28], quality and consistent progress updates [23], [26], [30], as well as the creators' backing history [33], creators' relationship with backers [34] and creators' personality traits [29], [35], play a crucial role in the success of the campaign. Mitra and Gilbert [27] showed that the language or specific phrases used by the project founders have a fundamental effect on attracting investors. These influential terms are primarily associated with 1) social identity, 2) reciprocity, 3) scarcity, 4) social proof, 5) liking, and 6) authority.

In crowdfunding, particularly on Kickstarter, projects incorporate textual components such as titles, abstracts, detailed descriptions, reward statements, and creator biographies, which are all presented in written form and encompass both objective and subjective elements. Wang et al. [41] found that (i) The subjective expression of titles, detailed descriptions, and biographies has a notably positive influence on fundraising outcomes. (ii) When crafting the detailed textual description, strategically placing objective content at the beginning of the narrative, followed by subjective statements, enhances the likelihood of achieving successful online fundraising.

The inability of creators to acquire a sufficient number of potential investors is one of the most common causes of project failure on Kickstarter [36]. An et al. [36] and Rakesh et al. [37] developed automatic recommendation systems that link projects with potential investors considering a variety of characteristics, including project-based, personal, location-based, and network-based characteristics. According to Gerber's study [25], the originality of project ideas, rewards, and motivation to aid the community can attract the appropriate investors. These findings provide valuable insights and guidance for designing project content for campaign developers.

Mollick et al. [38] found that the wisdom of the crowd appears to be comparable to that of professionals when it comes to determining whether or not to support a project. Lynn et al. [39] found that strangers in crowdfunding communities on Twitter play a direct role in the dissemination of knowledge and investment on the platform. Kim et al. [40] have focused on identifying the types of investors and the influential ones, finding out what types of investors often influence others in making investment decisions - product specialists and market specialists.

B. DECEPTION AND FRAUD DETECTION

As the online world has become a gold mine for deceivers and scammers, there has been much work on understanding and identifying liars, deceivers, and fraudsters [42], [43], by analyzing financial statement scams [44], deceptive emails [45], deception in online dating profiles [46], and various types of scammers on a popular online dating site [47]. These studies have found that deceivers leave behind traits or footprints of themselves in their content, behavior, language, as well as networking or content propagation patterns, when they attempt to deceive. In particular, text content manipulation, as we frequently observe with the falsification of information on social media, is one of the most simple, prevalent and inexpensive methods of deception. Success in deception is frequently high because of the absence of expertise and resources for fact-checking, transparency, as well as accountability. The linguistic method for spotting deception shows that the unconscious use of specific types of words can reveal the emotions and mental state of those who are lying, as word choices in communication can uncover various psychological and social aspects of individuals [49]. As a result, language analysis has been utilized to uncover fraud and deceit, including false identities [46], deceitful financial statements [44], and lying messages within businesses [45]. In computer-mediated text-based interactions, linguistic clues such as word count, pronouns, emotion words, and exclusivity words have been shown to be particularly beneficial in detecting fraud [43], [48], [50]. Due to their lack of experience and truth on what they are trying to falsely describe or discuss, and their need to avoid contradictions in their statements being made, deceivers have difficulty writing and therefore provide fewer details than truth-tellers [48], [51], [52]. The usage of first-person pronouns is known to show a person's ownership over their remark. Hence, deceivers typically avoid using it to distance themselves from their statements [50]. Creating a story takes a lot of cognitive resources; consequently, liars may avoid using words with cognitive processes (e.g., ''without'', ''except'', and ''but'') while lying [53].

Our previous work was one of the earliest studies that empirically showed the feasibility of detecting fraudulent crowdfunding projects using linguistic features, where we demonstrated that scammers intentionally try to deceive people by providing less information and writing more carefully and less casually [54]. In online debt crowdfunding (i.e., peer-to-peer lending), Gao and Lin [12] found that well-established characteristics of creditworthiness, such as readability, objectivity, signs of negativity, and deception, are meaningfully associated with loan repayment. In particular, a higher proportion of deception predictors (i.e., more spelling and grammar errors and a lack of objective, spatial, and temporal information) in a loan application are typically associated with an increased probability of default. Siering et al. [13] found that the use of linguistic and content-based cues resulted in a classification accuracy of 79.7% in detecting fraudulent crowdfunding campaigns. Cumming et al. [14] examined the correlation between various factors and the detection capability against crowdfunding frauds. They found that the success of detecting fraud was correlated to factors such as the details of the campaign description (i.e, readability, funding period duration), the background of the campaign creators, the availability of social information, and the characteristics of the campaign, such as the types of rewards offered. Lee et al. [55] found 17 features of content and creators of fraudulent crowdfunding campaigns, with which they built a fraud detection model with 87.3% accuracy.

The strengths of our paper over previous works are: we (i) use a broad range of traits (90 in total, to begin with) of crowdfunding campaigns inspired by prior related work on crowdfunding and lie or deception detection, and then (ii) propose to adopt and combine three well-known multiple feature selection methods, in order to find out more common and representative characteristics of fraudulent campaigns. (iii) As a result, we successfully identified a smaller set of 10 key representative features of frauds, three of which are new, original findings, yet capable of achieving classification performance (82.04% accuracy) comparable to those of the previous literature. (iv) We employ SHAP to further analyze the fraudulent classification model, providing an explanation of the importance of each feature.

III. METHODOLOGY

This section describes our methodology, including the dataset, and the set of features we propose to use, as well as their background hypotheses.

A. DATASET

Our dataset consists of publicly available data collected from Kickstarter [56]. A typical crowdfunding campaign in this dataset includes the following elements: 1) *Campaign* section, where the project creator presents and elaborates on their idea, often with supporting visual materials like videos and still images, 2) *Updates* section, which provides progress updates on the project, 3) *Comments* section, where backers and creators can communicate and discuss thoughts and feedback, 4) *Community* section, displaying information on the top 10 cities and countries where backers are located and the number of first-time and returning backers, and 5) *FAQ* section, which provides answers to frequently asked questions from backers.

Due to the lack of publicly available ground truth data of fraudulent crowdfunding campaigns, we collected hundreds of accused campaigns from various public forums such as Kickscammed.com [8], Reddit.com [7], and the Facebook page ''Crowdfunding Projects that Never Delivered'' [9]. These campaigns raised approximately \$11.5 million from 175,260 backers. A comprehensive review of all comments and updates for each project was conducted, spanning a minimum of one year after launch, up to five years for the oldest projects, to reduce the risk of incorrect or baseless allegations. From this, the list was narrowed down to 27 well-known fraudulent cases and 75 highly suspected frauds, for a total of 102 cases, based on the following criteria: (i) there are no signs that any backers have received project outcomes even after the promised delivery date, though it is still possible that someone received the delivery yet have not left any such messages at anywhere we have checked. (ii) There is no evidence that the allegation has been resolved elsewhere, including in the *Comments* section or public forums. When a project meets conditions (i) and (ii), it is added to our list of (highly) suspected frauds. (iii) In addition, if a campaign has been widely criticized as a fraud via press media coverage, such as Forbes.com, CNNMoney.com, etc., it is labeled as a well-known fraudulent case. We also collected data from 149 non-fraudulent campaigns from successfully delivered projects, based on their content in *Updates* and *Comments*. These successfully delivered projects were identified from various public forums such as the Facebook page ''The marketplace for successfully crowdfunded projects'' [57], and CNNMoney [58]. To collect all the data, we developed a crawler using the widely used Selenium library [59], enabling systematic crawling and downloading of publicly available data from Kickstarter.

We admit that our dataset is still research-grade; however, it comprises one hundred thoroughly examined cases, comparable in size to those used in previous research [13], [14]. For the protection of privacy information, we have taken measures to anonymize social information. Specifically, we have only verified the existence of a Facebook ID and any external links, without storing any identification information. Sensitive information from the *Campaign*, *Updates*, and *Comments* sections, such as email address, phone number, is anonymized by tokenization, using specific tokens such as ''<email>'' and ''<phone>''. Therefore, all direct identifiers of users have been removed, as well as social information.

B. FEATURE SET

Our feature set consists of i) campaign information (e.g., goal amount, # of backers, etc.), ii) profile and behavior features of campaign creators (e.g., backing history, comments, etc.), iii) social traits, and (iv) linguistic features (e.g., typo ratio, word counts, readability, etc.), extracted from *Campaign*, *Updates*, and *Comments* sections. We preprocessed the raw text data using word-level tokenization and lemmatization with NLTK [60], converting all text to lowercase. Additionally, we conducted part-of-speech tagging using the Stanford POS Tagger [61].

1) CAMPAIGN INFORMATION

To launch a campaign, creators are required to set a goal amount. To be successful, a campaign must achieve the target amount within the fundraising period, typically for a month or two. Regardless of the target goal amount, backers can continue to pledge money during the funding period. The currency used is the dollar; for projects using other currencies, amounts are converted to dollars based on the exchange rate at the project's launch date.

- Goal: Funding goal set by the campaign creator.
- Pledged: The total amount raised by backers.

In addition to the goal and pledged amount, we adopted using the total number of backers, the total number of reward backers striving to get promised rewards, and the total number of comments from backers. These features reflect the interest of the backers and their participation in a campaign.

- #_backers: The total number of backers (#_reward_backers + #_people who have pledged without asking for a reward).
- #_reward_backers: The number of backers who pledged money to get some reward.
- #_backers_comments: The number of comments from backers.

We also include the total number of videos and images included in *Campaign* and *Updates*; as uploading fake or manipulated videos or images may require much more care and costs than making real, authentic ones, fraudsters are less likely to use them. In 2012, there was a case where Kickstarter pulled the plug on the project named *"eye3"* [62] Affordable Flying Robot after people had noticed that the uploaded video of the project page was actually photoshopped images of an Xaircraft drone, which was already selling then.

- Videos: The total number of videos
- Images: The total number of images

2) BEHAVIORAL FEATURES OF CREATORS

Koch [63] found that creators who have backed others' projects have higher chances of success in fundraising. Wessel et al. have found that the creator can share their project on Facebook, and the act of trying to get funding from people by manipulating and increasing the likes of shared articles has a negative impact on the success of the project [64]. In the same context, we propose to use features capturing behavioral history of campaign creators; the number of comments left by the campaign creator on other campaigns before or after the launch of their own, the number of created campaigns, and the number of backed campaigns (i.e., # of campaigns the creator has pledged some amount).

- #_backed_campaigns: The number of campaigns the creator has pledged some amount of money.
- #_created_campaigns: The number of campaigns launched by the creator so far.
- #_before_comments: The number of comments left on other campaigns by the creator, before the launch of their own.
- #_after_comments: The number of comments left on other campaigns by the creator, after the launch of their own.

Studies have shown that entrepreneurs with greater potential, creativity, and (pro)activity have higher chances of leading a successful business [65]. To measure how (pro)active the creator of a project is in communicating with backers and reporting the progress to them, we propose to use the following features:

- #_public_updates: The number of publicly available *Updates*.
- #_updates: The total number of (public + for backers-only) *Updates*.
- #_comments: The total number of comments from the creator on the given campaign.

3) SOCIAL TRAITS

The success of crowdfunding [28] and the funding of new venture companies [66] have been reported to be related to the size of the creator's social network. Previous research has shown that information related to social networks has a significant influence on consumer decisions, e.g., movie box office revenue [67], book sales on Amazon [68]. Wessel et al. [14] and Lee et al. [55] claimed that people with fraudulent behavior tend to be reluctant to expose themselves on social media, as they intentionally avoid providing any information or clue to minimize the risks of being held accountable for their crimes by the public. Inspired by these studies, we add the following traits to our feature set.

- #_external_links: The number of external links or websites (e.g., LinkedIn, Twitter).
- Facebook_ID: Presence of Facebook ID.

4) LINGUISTIC CUES

Zuckerman and Depaulo [69] conceptualized lying as a more cognitively complex task than truth-telling. Newman et al. [53] found that liars can experience much more cognitive complexity. Because liars need a lot of imagination and effort when lying, there can be a greater cognitive load than those who tell the truth. Under a higher cognitive load, fraudsters are more likely to use fewer words, verbs, and sentences, with less diversity and redundancy in writing. We hypothesize that:

H1: Fraudsters experience greater cognitive load.

- Quantity: The number of words, verbs, and sentences
- Diversity
  - Lexical diversity: Percentage of unique words (the number of different words / the number of words)
  - Redundancy: The number of function words / The number of sentences
- Typo ratio: The number of misspelled words / The number of words

Launching campaigns and making progress updates available to backers, creators would often need to provide information on their own, as well as when and where they have been to, who they meet or work with, etc. We hypothesize that fraudsters may not afford to reveal their data or information in order to avoid or reduce the chances of public scrutiny. We hypothesize that:

H2: Fraudsters try to hide their personal or NER information.

- Email: The number of times the creator mentions email addresses
- Named-entity recognition words [71]: The number of Named-entity recognition words (i.e., names of person, location, and organization, like James, London, and Facebook)

Lying is related to psychologically negative emotions such as guilt, anger, and anxiety. When people lie, they feel discomfort and guilt [51], [52], [69], [72]. Newman [53] found that liars used more negative emotion words. We use the Linguistic Inquiry Word Count (LIWC) dictionary [70] to measure emotion words. We hypothesize that:

H3: Fraudsters use less positive and more negative emotion words.

- positive emotion: The number of positive emotion words (e.g., "love", "nice", "good")
- negative emotion: The number of negative emotion words (e.g., "hate", "worthless", "sorry")

The method of measuring psychological distance in content is also often used to find the characteristics of liars. Psychological distance is a subjective experience of reality that the following things are close or far away from the self, here, and now [73]: (i) temporal (e.g., my first year of university), (ii) spatial (e.g., my house, universe), (iii) social distance (e.g., my best friends, co-workers), and (iv) hypothetical alternatives to reality, what would or might have been but did not happen (e.g., If I was born 100 years ago). Wiener and Mehrabian [74] reported that liars were more "non-immediate" than truth-tellers, referring to themselves less often in their stories. Newman et al. [53] observed that liars use first-person singular pronouns less frequently than truth-tellers, attempting to disassociate themselves from lies by projecting less of themselves on their contents. It has been known that using first-person singular pronouns involves taking ownership of a statement, by which they imply that they are being honest with themselves [53]. We predict that fraudulent campaign creators would use fewer singular first-person pronouns, but more plural, second, and third-person pronouns, instead. We hypothesize that:

H4: Fraudsters avoid using first-person pronouns, being reluctant to take ownership of their content.

- First-person singular: The number of first-person singular pronouns
- First-person plural: The number of first-person plural pronouns
- Second-person: The number of second-person pronouns
- Third-person: The number of third-person pronouns

Lying is cognitively difficult because it manipulates information by costly creating false one, which requires a lot of cognitive resources [51]. As a result, liars often avoid the use of words with high cognitive processes, which are known to place heavy demands on cognitive resources while lying [53]. Cognitive process words are also measured using the LIWC dictionary. We hypothesize that:

H5: Fraudsters use fewer cognitive process words.

- Exclusive words: The number of exclusive words. (e.g., "except", "without", "but")
- Insight: The number of insight words. (e.g., "think", "know", "consider")
- Causal: The number of causal words. (e.g., "because", "effect", "hence")

Readability tests are used to evaluate the readability of text by counting syllables, words, and sentences. Moffitt and Burns [75] showed that deceptive financial reports have a higher word complexity (e.g., including more qualifying conjunctions), which means low readability. The higher readability score means, the easier it is to read [15]. We hypothesize that:

H6: Fraudulent campaigns have a lower (text) readability.

- ARI: The Automated Readability Index [15]
- CL: The Coleman-Liau index [16]
- GF: Gunning Fog index [17]
- FKGL: Flesch-Kincaid Grade level [18]
- FRES: Flesch-Reading Ease score [18]

Readability metrics are described in detail in the Appendix A.

C. FEATURE SELECTION

Feature selection aims to extract representative variables for effective prediction from a given dataset. It has the advantages of overfitting suppression, increased accuracy, as well as decreased training time [20]. In this paper, we propose to use and combine multiple feature selection methods to extract a set of more representative and reliable features by leveraging their complementary strengths. Specifically, we propose to extract and use features commonly selected by (i.e., the intersection of) the three well-known feature selection methods: CFS, PCC, and IG.

CFS evaluates feature subsets based on the principle that good feature subsets contain features highly correlated with the target variable (i.e., fraudulent campaign) but uncorrelated with each other [19]. By considering the relationships between features, it minimizes redundancy while retaining relevant features, thus improving model performance. PCC evaluates individual features by measuring their linear correlation with the target variable, prioritizing features based on their relevance [20]. IG is an entropy-based measure that quantifies the reduction in uncertainty regarding the target variable by comparing its entropy before and after partitioning the data based on a specific feature [19]. By integrating these three methods, which use distinct approaches to evaluate features, we achieve a complementary selection process that ensures the inclusion of relevant, informative, and non-redundant features, resulting in a more robust and representative feature set for prediction [21]. These methods are used using default parameters on Weka [77].

1) CORRELATION-BASED FEATURE SELECTION (CFS) [19]

A well-known method developed to evaluate the importance of different subsets of features within a dataset, CFS measures both the predictive capability of each feature in relation to the target variable and the degree of overlap (or redundancy) among the features. Initially, CFS calculates correlation matrices to determine the strength of the relationship between each feature and the target variable, as well as the level of correlation among the features themselves. Based on these calculations, CFS identifies the most effective subset of features by employing a best-first search strategy [79]. The ideal subset comprises features that are strongly correlated with the target outcome but have minimal correlation with each other, ensuring a balance between strong predictive power and low redundancy. The effectiveness of a feature subset is evaluated using the following equation:

$$M_s = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \tag{1}$$

In this equation, M_s denotes the merit of the selected feature subset, where k represents the total number of features under consideration. The term $\overline{r_{cf}}$ is the average correlation between each feature and the class(i.e., target variable), highlighting the predictive strength of features. $\overline{r_{ff}}$ represents the average correlation between every pair of features, indicating their redundancy.

2) PEARSON CORRELATION COEFFICIENT (PCC) [20]

Feature selection based on PCC evaluates the linear relationship between individual features and the target variable within a dataset using the Pearson correlation coefficient. This method calculates the Pearson correlation for each feature in relation to the target variable, quantifying the strength and orientation of the linear association. The correlation coefficient values range between -1 and $+1$, where values closer to $+1$ indicate a strong positive linear relationship, values closer to -1 signify a strong negative linear relationship, and value of 0 denotes no linear correlation. Integrating a ranking method with Pearson correlation allows for evaluating features based on their relevance and impact. This prioritizes features based on the absolute value of their correlation coefficients, emphasizing their influence on the target irrespective of the direction of the relationship.

The significance of both positive and negative correlations is considered, with positive correlations indicating that as the feature value increases, the target variable also increases(i.e., the probability of the target variable being classified into a particular class increases in the context of binary outcomes), and negative correlations indicate that as the feature value increases, the target variable decreases(i.e., the probability of the target variable being classified into an opposite class increases). This approach ensures that features strongly associated with the target variable, whether positively or negatively, are highlighted. Features are ranked and selected based on their absolute correlation values, prioritizing those with the strongest linear relationship to the target variable, as illustrated by the equation:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{2}$$

In this formula, r_{xy} represents the Pearson correlation coefficient between a feature and the target variable, where x_i and y_i denote the values of the feature and the target, respectively. $\overline{x}$ and $\overline{y}$ indicate the mean values of the feature and target. This method facilitates the identification and prioritization of features based on their direct linear correlation with the target, facilitating the selection of a highly relevant subset of features for predictive modeling by considering both the magnitude and direction of their relationships.

3) INFORMATION GAIN (IG) [19]

IG measures the reduction in uncertainty about the target variable due to the use of a feature. This metric aids in identifying the most informative features for prediction. It is calculated based on the difference in entropy before and after the feature is observed. Entropy, a foundational concept in information theory, measures the unpredictability or randomness within a system. Mathematically, the entropy of a target variable Y, prior to the observation of any feature, is defined as:

$$H(Y) = -\sum_{y \in Y} p(y)\log_2(p(y)) \tag{3}$$

where $p(y)$ denotes the probability of occurrence of each possible value of Y. When a new feature X is introduced, the conditional entropy of Y given X is computed as:

$$H(Y|X) = -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x)\log_2(p(y|x)) \tag{4}$$

$p(x)$ and $p(y|x)$ represent the probability of observing X and the conditional probability of Y given X, respectively. IG represents the reduction in entropy due to the observation of X, is calculated as:

$$IG(Y|X) = H(Y) - H(Y|X) = H(X) - H(X|Y) \tag{5}$$

IG is a symmetric metric, meaning that the information gained about Y from observing X is identical to the information gained about X from observing Y.

D. PERFORMANCE METRICS

The following elements are used in performance metrics.

1) True Positive(TP): Correctly classified as fraudulent projects
2) True Negative(TN): Correctly classified as non-fraudulent projects
3) False Positive(FP): Non-fraudulent projects falsely labeled as fraudulent projects
4) False Negative(FN): Fraudulent projects falsely labeled as non-fraudulent projects

We employ four measures to evaluate the classification performance of our fraud detector: (overall) accuracy, AUC, precision, and recall.

- (Overall) accuracy: measures the ratio of correctly classified fraud or non-fraud projects to the total number of projects in the dataset. We applied this metric to measure the accuracy of a classifier on our entire dataset. The following equation shows the calculation for accuracy:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

- AUC [76]: The Area Under the Receiver Operating Characteristic (AUC) is calculated as the area under the ROC curve. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at different threshold values, giving a visual representation of how the model performs at different classification thresholds. The AUC score ranges from 0 to 1, with a score of 1 indicating a perfect classifier that can correctly distinguish between positive and negative cases with 100% accuracy.

In addition to these two metrics, the following metrics are also used to evaluate per-class detection performance, particularly for frauds:

- Precision: measures the proportion of correctly classified fraudulent projects among all the projects that the model predicted as fraudulent projects. This indicates how precise the model's positive predictions are. The following equation shows the calculation for precision:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

- Recall: measures the proportion of correctly classified fraudulent projects among all the actual fraudulent projects in our dataset. This indicates how many fraudulent projects the model correctly identified. The following equation shows the calculation for recall:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

IV. RESULTS

This section initially focuses on extracting useful and discriminative characteristics of fraudulent crowdfunding projects. We then examine the feasibility of building an accurate classification model for fraud detection.

A. DISTINGUISHING CHARACTERISTICS OF FRAUDULENT CAMPAIGNS

Table 1 presents the results of features commonly selected from the three feature selection algorithms, as well as each feature selection algorithm. The CFS algorithm identified 19 features from the total of 90 input features, and PCC and IG selected the top 20 (by rank) features that contributed the most to classification based on rank scores. We successfully identify 10 commonly selected key features of fraudulent campaigns, as shown in Table 1a, three of which are our own new, original contributions. We obtained 3, 3, and 4 features from Campaign & Creator information, *Comments*, and *Updates* section, respectively, none from the *Campaign* section. Table 2 shows the classification performance when each category of commonly selected features was used exclusively, with Random Forest algorithm. Features extracted from Campaign & Creator information and *Comments* section are good predictors for fraud detection, achieving 67.3% and 65.7% accuracy, respectively. Our model achieves up to 82.04% accuracy with 10 commonly selected features and the Random Forest algorithm.

We now examine and discuss key selected distinguishing characteristics of fraudulent campaigns. To verify our null hypothesis that the mean of fraudulent and non-fraudulent projects is equal, we adopt the un-paired t-test. This method proves to be efficient and reliable, even in scenarios where the dataset size is small [80], [81]. The statistical significance of the t-test refers to whether the difference between "fraud" and "non-fraud" averages reflects a real difference in the population from which the groups were sampled. Tables 6, and 7 in Appendix B present descriptive statistics for fraudulent and non-fraudulent campaigns in three sections (*Campaign*, *Updates*, and *Comments*, respectively).

1) FEATURES FROM CAMPAIGN & CREATOR INFORMATION AND SOCIAL TRAITS

The number of backed projects by the creator, and the number of external links and websites, are good predictors for detecting frauds, which is in line with [14] and [55]. Additionally, as an original finding of ours, the number of videos in *Updates* section is also a good predictor. Our classifier correctly classified fraudulent projects with 67.33% accuracy using these three variables, as shown in Table 2.

As shown in Table 3, videos are used 3.3 times more frequently in *Updates* section of non-fraudulent campaigns than fraudulent ones (M_F: 0.032, M_N: 0.108, $p < 0.001$). M_F and M_N are the mean values for frauds and non-frauds, respectively. This result indicates that fraudsters find it more difficult to produce and upload video content, particularly

TABLE 1. Features commonly selected by the three algorithms, CFS, PCC, and IG.

(a) Commonly selected features by all three algorithms.

Algorithms	Campaign & Creator info.	Updates	Comments
Commonly Selected Features	#_backed_campaigns #_external_links Video / #_public_updates	Typo ratio Email / #_public_updates Location / #_public_updates First-person plural / #_public_updates	Email / #_comments Insight / #_comments Causal / #_comments

(b) Other features selected by each of the algorithms.

Algorithms	Campaign & Creator info.	Updates	Comments
CFS	Pledged #_backers #_reward_backers #_before_comments	#_sentences / #_public_updates Person / #_public_updates	Redundancy First-person singular / #_comments FRES
PCC	Facebook_ID	Second-person / #_public_updates CL	#_total_words / #_comments #_verbs / #_comments Redundancy Typo ratio Exclusive / #_comments GF FKGL
IG	Pledged #_backers #_reward_backers	#_total_words / #_public_updates #_sentences / #_public_updates #_verbs / #_public_updates Person / #_public_updates Positive_emotion / #_public_updates Second-person / #_public_updates ARI	-

TABLE 2. Classification performance of our model built with each category of commonly selected features using Random Forest (Precision and Recall on frauds).

Feature	Precision	Recall	Accuracy	AUC
Campaign & Creator info.	60.70%	57.81%	67.33%	0.658
Campaign	-	-	-	-
Updates	57.51%	49.81%	64.10%	0.618
Comments	58.65%	53.81%	65.73%	0.638
Total	82.19%	71.27%	82.04%	0.802

in *Updates* section. We also found that non-fraudsters invest more actively than fraudsters on the platform. #_backed_campaigns of non-fraudsters are 2.67 times higher than fraudsters (M_F: 8.568, M_N: 22.959, $p < 0.001$), i.e., non-fraudsters tend to invest more often in other campaigns of their interest. This implies that creators with a more active backing experience on a crowdfunding platform are less likely to set up a fraudulent campaign. Non-fraud campaigns contain an average of 2.438 external links or websites, whereas frauds contain an average of 1.568, which suggests that fraudsters are more reluctant to reveal or provide further information in addition to the campaign contents, particularly their SNS and blog pages.

2) FEATURES FROM *CAMPAIGN*, *UPDATES*, AND *COMMENTS* SECTIONS

As shown in Table 1, no features in *Campaign* section were selected during our feature selection process. Looking at the results of the t-test (as shown in Table 6) for the *Campaign* section, except the case of typo ratio, the level of significance of all features is higher than 0.05, indicating *Campaign* content is not that helpful in characterizing and detecting fraudulent campaigns.

As the number of updates and comments varied across different campaigns, we normalized all *Updates* and *Comments* features before performing our experiments. Thus, except for proportional features like typo ratio and redundancy, we consider all those features per each update (e.g., the average number of words per update) and per each comment (e.g., the average number of words per comment).

In *Updates* section, we found that fraudsters make fewer typographical errors than non-fraudsters (M_F: 0.045, M_N: 0.060, $p < 0.001$). Similar patterns were observed in *Campaign* and *Comments* sections as well, which all consistently suggest that fraudulent campaign creators seem to write their text content more carefully, deliberately trying to deceive backers and potential investors.

We also found that non-fraudsters referred to email information 3.3 and 2.2 times more than non-fraudsters in *Updates* (M_F: 0.049, M_N: 0.162, $p < 0.001$) and *Comments* (M_F: 0.056, M_N: 0.124, $p < 0.001$), respectively. This result can be interpreted as a reluctance of fraudsters to disclose their contact information, which is consistent with previous literature [55]. Non-fraudsters used twice as many location words (M_F: 0.740, M_N: 1.416, $p < 0.001$), e.g., London, Seoul, Paris. They also tend to use more names of organizations and people than fraudsters, which support Hypothesis 2, that fraudsters try to hide their personal or Named-Entity Recognition information. Due to the dilemma

TABLE 3. Campaign & Creator information and Social Traits, the mean and standard deviation of fraud and non-fraud projects, and t-tests for the difference. Significance: *** p < 0.001, ** p < 0.01, * p < 0.05. If p-value is less than 0.05, it becomes more significant.

Category	Features	Fraud Mean	Fraud SD	Non-fraud Mean	Non-fraud SD	Difference p-value
Campaign info.	Goal	38,816.403	64,836.050	68,204.008	102,818.740	**
	Pledged	129,242.499	393,104.756	398,477.524	919,957.862	***
	#_backers	1,317.715	2,484.527	4,226.335	9,624.865	***
	#_reward_backers	1,296.225	2,470.260	4,162.630	9,542.621	***
	#_backers_comments	964.137	2,383,748	1,252.469	3,148.618	0.434
	#_campaign_videos	0.382	0.833	0.771	1.555	*
	#_campaign_images	16.901	16.623	13.550	11.216	0.057
	#_updates_videos / #_public_updates	0.032	0.116	0.108	0.201	***
	#_updates_images / #_public_updates	0.781	1.088	1.024	1.252	0.113
Creator info.	#_backed_campaigns	8.568	15.287	22.959	34.704	***
	#_created_campaigns	1.725	1.536	2.409	2.691	*
	#_before_comments	14.862	56.169	8.932	54.196	0.402
	#_after_comments	10.235	41.913	29.818	102.802	0.069
	#_updates	23.196	20.964	31.342	25.139	**
	#_public_updates	16.735	18.618	24.691	24.174	**
	#_comments	73.598	134.932	94.932	161.748	0.274
Social Traits	#_external_links	1.568	1.582	2.483	1.587	***
	Facebook_ID	0.343	-	0.550	-	**

of the deceiver, deceivers try to conceal and avoid revealing or releasing their information, such as names of location and person [82], [83]. Liars tend to provide less or unverifiable information instead of providing verifiable information. When trying to describe a venue or location they have not been to, it is very difficult for them to create spatial details because they have to create imaginative and fictitious writing [84], [85].

Notably, non-fraudsters use 1.43 times more first-person plural pronouns in *Updates* (M_F: 5.186, M_N: 7.44, $p < 0.001$), which is inconsistent with the previous literature on computer-mediated interactive communication [50], not supporting Hypothesis 4. It is possibly because those non-fraudulent crowdfunding campaign creators typically work as a team, so they naturally often refer to themselves using first-person plural pronouns, as supported by the following literatures; According to [87], only 28% of approximately 10,000 crowdfunding projects are carried out individually, whereas 72% campaigns were carried out in teams. Franke et al. [88] found that backers of campaigns or ventures tend to be more confident when a campaign or venture is carried out by a team instead of an individual. Beier and Wagner [37] also found that specific campaign characteristics like team size and national proximity leverage fundraising success.

In *Comments* section, as exemplified in Table 4, fraudsters turned out to use more cognitive process words such as insight words (i.e., "think", "know", with M_F: 1.563, M_N: 0.937, $p < 0.001$) and causal words (i.e., "because", "hence", with M_F: 1.453, M_N: 0.892, $p < 0.001$), which are our original findings. Interestingly, the results are in line with [89], whereas inconsistent with [53] and our own results observed in *Updates* section. We speculate that this inconsistency is due to the content type and communication mode used. In the case of [53] and our *Updates* section data, the contents are mainly filled or written in a one-way, non-interactive, presentational mode, in which Newman et al. found that scammers and liars tend to use fewer cognitive process words, because they have to create or report a whole fake story, which demands a lot of cognitive resources. In contrast, our *Comments* section data mostly consists of interactive communications between creators and backers, which is consistent with Ho et al.'s study [89] where they also found that deceivers used more words of insight (i.e., "think", "know"), when tested with their own dataset of interactive dialogues collected in an online game. To get a glimpse of how and in what context fraudsters often use those words, we present typical examples of such comments in Table 4. We see that words of insight and causation are often used by fraudsters to (i) make false excuses to justify the lack of progress or unpreparedness, and delay of the reward delivery schedules, as well as (ii) present subjective beliefs or claims rather than objective, verifiable facts, feeling extra fear from worrying about being caught than truth-tellers.

TABLE 4. Example comments from fraudsters using insight and causal words examples.

"@LT - We are choosing not to set out specific dates because we cannot guarantee consistent production"
"Hi Everyone, We delayed our update because as of last Thursday, nothing had changed."
"we will let you know"
"Please let us know if you have any more questions."
"I think we're not doing too bad"
"As a result, currently there is absolutely NO LEGAL ACTION POSSIBLE on us."
"I want to make sure that your first experience with Dysis is a blast, even though it will just be a Prototype!"

B. CLASSIFICATION PERFORMANCE EVALUATION

We next measure the discriminative power of the selected 10 features of fraudulent crowdfunding campaigns, by

TABLE 5. Classification performance of six machine learning algorithms (10-fold cross-validation).

Algorithm	Precision	Recall	Accuracy	AUC
Random Forest	82.19%	71.27%	82.04%	0.802
Logistic Regression	87.14%	60.63%	80.44%	0.773
SVM	82.01%	62.63%	79.26%	0.766
KNN(k=9)	77.28%	60.63%	76.86%	0.742
J48 Decision Tree	61.96%	65.54%	68.50%	0.680
Naive Bayes	57.03%	80.36%	67.30%	0.693

building and evaluating the classification performance of six well-known machine learning algorithms: Naive Bayes (NB), Support Vector Machines (SVM), J48 Decision Tree (J48), Random Forest (RF), k-Nearest Neighbors (k-NN), and Logistic Regression (LR). In this study, the classification models are implemented and evaluated using the free machine learning software Weka (with default parameters) [77]. The evaluation is conducted using 10-fold cross-validation with 90% of the data used for training and 10% for testing. To create a classifier with the best accuracy, we use the 10 commonly selected features by the three multiple feature selection methods, shown in Table 1. As shown in Table 5, RF performs best for accurate classification of fraudulent and non-fraudulent campaigns, with an overall accuracy of 82.04%, 82.19% precision, 71.27% recall, and 0.802 AUC, followed by LR (80.44% accuracy and 0.773 AUC), SVM (79.26% accuracy and 0.766 AUC), k-NN (76.86% accuracy and 0.742 AUC), J48 (68.5% accuracy and 0.68 AUC), and then NB (67.3% accuracy and 0.693 AUC). Our proposed method is capable of achieving classification performance comparable to those of the previous literature, with an accuracy of 82.04% [12], [13], [55]. The precision and recall analysis of RF and LR, which are the top-performing algorithms, shows that precision scores are higher than recall by approximately 11-27%. LR achieves a precision of 87.14%, demonstrating a high level of accuracy in classifying fraudulent campaigns. This indicates that the models are effective at minimizing false positives, which is crucial for maintaining user trust, but the relatively lower recall suggests that some fraudulent campaigns may go undetected. In contrast, The NB classifier achieved the highest recall rate of 80.36% among all algorithms, highlighting its strong capability to detect fraudulent campaigns, particularly in scenarios where minimizing false negatives is critical. However, this high recall is offset by its lowest precision score of 57.03%, indicating a significant trade-off. Considering this trade-off, RF emerges as a more effective choice for detecting fraudulent campaigns, providing a well-balanced approach that jointly considers both precision and recall. Our proposed approach successfully identified a smaller set of 10 key characteristics of frauds, with which our model is capable of achieving classification performance comparable to the previous literatures. Besides, three of those selected features – (i) fewer videos in *Updates*, (ii) more use of insight and (iii) causal words in *Comments* – are our own new, original findings.

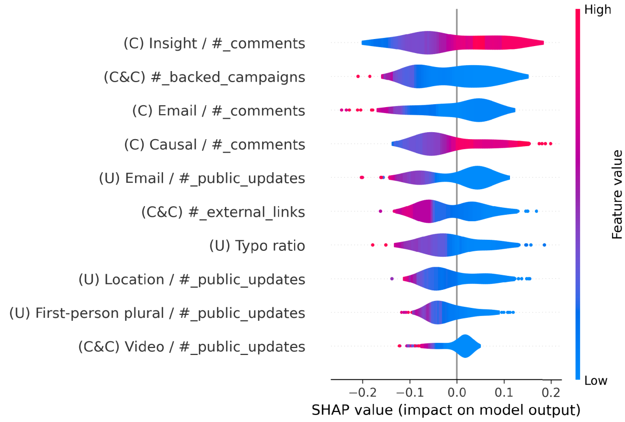

FIGURE 1. SHAP summary plot showing the contribution of features to the fraudulent campaign detection model (Random Forest, the best-performing model). (C&C): Campaign & Creator information, (U): Updates section, (C): Comments section.

C. MODEL EXPLAINABILITY

We examine how features contribute to the fraudulent campaign detection model using SHAP [90]. SHAP values quantify each feature's contribution to the model's predictions, providing a mathematically consistent explanation of their impact on the decision-making process. This approach helps elucidate the model's reasoning behind classifying fraudulent campaigns. The SHAP summary plot, as shown in Figure 1, provides a visual representation of the importance and distribution of each feature's impact on the model's predictions. In the SHAP summary plot, each row represents a feature, and each dot corresponds to a specific data sample. The x-axis denotes the SHAP value (impact on the model's output), while the color of each dot represents the magnitude of the feature's value in the respective sample (e.g., red for high values and blue for low values). Positive SHAP values indicate that the corresponding feature increases the predicted probability of fraudulent campaigns, while negative SHAP values decrease it. Features at the top of the plot are the most influential in the model's predictions.

Figure 1 presents the SHAP analysis conducted on the Random Forest model, which achieved the highest accuracy in Table 5. As shown in Figure 1, insight and causal words from the *Comments* section, two of the three newly identified features in this study, rank first and fourth, respectively, in their positive impact on the fraudulent campaign detection model. When the values of these features are high (represented by red dots), the likelihood of a campaign being classified as fraudulent significantly increases. This observation strongly reinforces the findings discussed earlier, where we noted that words of insight and causation are often used by fraudsters to (i) make false excuses to justify the lack of progress or unpreparedness, and delay of reward delivery schedules, as well as (ii) present subjective beliefs or claims rather than objective, verifiable facts, possibly due to heightened fear from worrying about being caught compared to truth-tellers. The number of backed projects by the creator shows that lower values (blue dots) increase the likelihood of a fraudulent campaigns, while higher values (red dots) reduce it, suggesting that creators with more crowdfunding investment experience are generally perceived as more trustworthy. The *Comments* section emerges as the most significant contributor to fraudulent campaign detection model, including insight and causal words, and email mentions, rank among the top four most impactful features. This underscores the critical role of linguistic patterns reflected in the *Comments* section content by project creators.

V. CONCLUSION

Crowdfunding has grown in popularity recently as a new method for startups and developing businesses to raise funds. As crowdfunding grows in popularity, there is also a significant risk of fraud. Despite growing concern over the growing threat of fraudulent crowdfunding projects, little is known about them, primarily due to a lack of measurement data collected from real-world datasets. We collected and analyzed hundreds of fraudulent campaigns from one of the most popular and publicly available crowdfunding sites, Kickstarter.com. We begin with 90 features inspired by previous related studies, contained in campaign information, campaign creators' profile and behavior information, and linguistic features that have been proved useful in deception detection. We strategically use and combine three well-known multiple feature selection methods, based on CFS, PCC, and IG, in order to identify representative features of fraudulent campaigns. We identified 10 commonly selected key features, three of which were our own original findings. Based on the 10 key features, our model classified fraudulent campaigns with 82.04% accuracy. To identify the most important features in the model, we used SHAP to demonstrate that insight and causal words, frequently used by fraudsters in the *Comments* section to make false excuses for delays, justify lack of progress, or present subjective beliefs rather than objective facts, ranked as the first and fourth most impactful features among the 10 commonly selected key features.

We admit that our work has limitations. There is no legal proof or evidence that the campaigns contained in our dataset are 100% absolute scams. We acknowledge that our dataset is still research-grade; however, it comprises over one hundred thoroughly reviewed fraudulent instances, making it comparable in size to those used in earlier studies. We use Kickstarter, one of the largest and most active crowdfunding platforms, but relying on a single platform limits the generalizability of our findings. Other platforms, such as Indiegogo or GoFundMe, may exhibit different patterns or characteristics of fraudulent behavior. Therefore, as part of our future research, we plan to expand the dataset to include multiple platforms to further enhance the generalizability of our results. We also aim to explore the use of deep learning algorithms such as BERT or GPT, which are based on transformers, to automatically learn useful linguistic features from raw text data and construct classification models. Furthermore, we propose possibilities for finding and analyzing useful words and sentences that aid in the classification of fraudulent projects using explainable algorithms.

APPENDIX A
READABILITY

- #_characters: The number of characters.
- #_words: The number of words.
- #_sentences: The number of sentences.
- #_complex: The number of complex words. Those with three or more syllables.
- #_syllables: The number of syllables

1. ARI: The Automated Readability Index is designed to gauge the understandability of a text. Representation of the US grade level. Index score ranges from 1 (Kindergarten) $\sim$ 14 (College). The lower the number, the easier it is to read.

$$\mathrm{ARI} = 4.71\frac{\#\_\mathrm{characters}}{\#\_\mathrm{words}} + 0.5\frac{\#\_\mathrm{words}}{\#\_\mathrm{sentences}} - 21.43 \tag{9}$$

2. CL: Coleman-Liau index is designed to gauge the understandability of a text. Representation of the US grade level. Index score ranges from $\leq$6 (6th grade) $\sim$ 14 (college sophomore). The lower the number, the easier it is to read.

$$\mathrm{CL} = 5.88\frac{\#\_\mathrm{characters}}{\#\_\mathrm{words}} - 29.59\frac{\#\_\mathrm{words}}{\#\_\mathrm{sentences}} - 15.8 \tag{10}$$

3. GF: Gunning Fog index is a test of English writing. The index estimates the years of formal education a person needs to understand the text on first reading. The index score ranges from $\leq$6 (6th grade) $\sim$ 17+ (college graduate). The lower the number, the easier it is to read.

$$\mathrm{GF} = 0.4[\frac{\#\_\mathrm{words}}{\#\_\mathrm{sentences}} + 100\frac{\#\_\mathrm{complex}}{\#\_\mathrm{words}}] \tag{11}$$

4. FKGL: Flesch-Kincaid Grade level is designed to indicate how difficult the passage is to understand in English. The score ranges from $\leq$ (6th grade) $\sim$ 14 (college sophomore).

$$\mathrm{FKGL} = 0.39\frac{\#\_\mathrm{words}}{\#\_\mathrm{sentences}} + 11.8\frac{\#\_\mathrm{syllables}}{\#\_\mathrm{words}} - 15.59 \tag{12}$$

5. FRES: Flesch-Reading Ease score is designed to indicate how difficult the passage is to understand in English. The score ranges from 0 (college graduate) $\sim$ 100 (5th grade). The lower the number, the harder it is to read.

$$\mathrm{FRES} = 206.835 - 1.015\frac{\#\_\mathrm{words}}{\#\_\mathrm{sentences}} - 84.6\frac{\#\_\mathrm{syllables}}{\#\_\mathrm{words}} \tag{13}$$

TABLE 6. The Campaign and Updates sections, the mean and standard deviation of fraud and non-fraud linguistic cues, and t-tests for the difference. Significance: *** p < 0.001, ** p < 0.01, * p < 0.05. If p-value is less than 0.05, it becomes more significant.

Section	Linguistic cues	Fraud Mean	Fraud SD	Non-fraud Mean	Non-fraud SD	Difference p-value
Campaign	H1: Fraudsters experience greater cognitive load.					
	Quantity					
	Word	1,079.666	675.904	1,110.248	754.720	0.742
	Verb	173.852	107.183	176.570	112.823	0.848
	Sentence	52.735	35.200	57.651	44.182	0.349
	Diversity					
	Lexical diversity	0.428	0.102	0.433	0.093	0.724
	Redundancy	5.100	2.365	4.619	1.620	0.056
	Typo ratio	0.063	0.023	0.070	0.026	*
	H2: Fraudsters try to hide their personal or NER information.					
	E-mail	0.068	0.290	0.140	0.419	0.132
	Named-entity recognition					
	Person	11.549	15.626	13.926	19.141	0.299
	Location	5.539	8.852	7.389	7.908	0.084
	Organization	11.568	13.640	12.248	19.649	0.762
	H3: Fraudsters use less positive and more negative emotion words.					
	Positive emotion words	40.637	27.651	40.993	27.472	0.919
	Negative emotion words	9.745	9.399	7.630	7.721	0.061
	H4: Fraudsters avoid using first-person pronouns, being reluctant to take ownership of their content.					
	Personal pronouns					
	First-person singular	2.323	5.364	1.791	4.070	0.373
	First-person plural	15.372	14.908	18.812	14.994	0.074
	Second-person	27.725	22.407	27.510	22.004	0.939
	Third-person	19.480	17.727	19.234	14.936	0.908
	H5: Fraudsters use fewer cognitive process words.					
	Exclusive words	23.117	16.331	24.355	19.652	0.600
	Insight	15.460	11.501	14.939	10.227	0.706
	Causal	29.078	20.725	28.067	17.241	0.685
	H6: Fraudulent campaigns have a lower (text) readability.					
	Automated Readability Index	12.710	5.156	11.904	4.002	0.165
	Coleman-Liau Index	10.929	2.525	10.930	2.082	0.996
	Gunning-Fog Score	13.979	4.395	13.248	3.120	0.124
	Flesch-Kincaid Grade Level	10.334	4.101	9.582	3.078	0.098
	Flesch-Reading Ease score	61.231	14.492	63.923	12.616	0.119
Updates	H1: Fraudsters experience greater cognitive load.					
	Quantity					
	Word / #_public_updates	215.963	164.308	255.310	190.951	0.091
	Verb / #_public_updates	38.861	28.947	46.260	34.743	0.077
	Sentence / #_public_updates	11.867	8.346	14.681	10.712	*
	Diversity					
	Lexical diversity	0.296	0.156	0.261	0.123	0.057
	Redundancy	4.096	1.473	4.019	0.908	0.638
	Typo ratio	0.045	0.021	0.060	0.022	***
	H2: Fraudsters try to hide their personal or NER information.					
	E-mail / #_public_updates	0.049	0.097	0.162	0.262	***
	Named-entity recognition					
	Person / #_public_updates	1.862	2.271	2.521	2.473	*
	Location / #_public_updates	0.740	0.881	1.416	1.602	***
	Organization / #_public_updates	1.593	2.118	1.831	1.740	0.331
	H3: Fraudsters use less positive and more negative emotion words.					
	Positive emotion words / #_public_updates	9.029	6.274	11.040	6.849	*
	Negative emotion words / #_public_updates	1.588	1.608	1.690	2.667	0.731
	H4: Fraudsters avoid using first-person pronouns, being reluctant to take ownership of their content.					
	Personal pronouns					
	First-person singular / #_public_updates	0.741	0.893	0.681	2.028	0.780
	First-person plural / #_public_updates	5.186	5.305	7.440	5.166	***
	Second-person / #_public_updates	5.075	3.792	7.117	4.638	***
	Third-person / #_public_updates	3.722	3.059	4.548	5.257	0.154
	H5: Fraudsters use fewer cognitive process words.					
	Exclusive words / #_public_updates	5.316	4.004	6.216	5.668	0.167
	Insight / #_public_updates	3.779	3.069	4.090	3.616	0.478
	Causal / #_public_updates	4.438	3.869	5.086	4.103	0.210
	H6: Fraudulent campaigns have a lower (text) readability.					
	Automated Readability Index	8.696	3.312	9.353	2.382	0.086
	Coleman-Liau Index	8.264	2.587	9.330	1.925	***
	Gunning-Fog Score	10.866	3.387	11.275	1.751	0.264
	Flesch-Kincaid Grade Level	7.046	2.652	7.191	1.787	0.630
	Flesch-Reading Ease score	71.539	19.833	76.864	8.414	*

TABLE 7. The comments section, the mean and standard deviation of fraud and non-fraud linguistic cues, and t-tests for the difference. Significance: *** p < 0.001, ** p < 0.01, * p < 0.05. If p-value is less than 0.05, it becomes more significant.

Section	Linguistic cues	Fraud Mean	Fraud SD	Non-fraud Mean	Non-fraud SD	Difference p-value
Comments	H1: Fraudsters experience greater cognitive load.					
	Quantity					
	Word / #_comments	68.111	52.651	51.205	27.146	**
	Verb / #_comments	13.171	10.014	9.666	5.223	**
	Sentence / #_comments	3.749	2.575	3.328	1.849	0.133
	Diversity					
	Lexical diversity	0.328	0.180	0.317	0.180	0.625
	Redundancy	4.147	1.804	3.432	1.195	***
	Typo ratio	0.057	0.021	0.067	0.025	**
	H2: Fraudsters try to hide their personal or NER information.					
	E-mail / #_comments	0.056	0.127	0.124	0.171	***
	Named-entity recognition					
	Person / #_comments	1.026	1.239	1.251	0.812	0.083
	Location / #_comments	0.147	0.221	0.183	0.277	0.281
	Organization / #_comments	0.258	0.353	0.200	0.224	0.144
	H3: Fraudsters use less positive and more negative emotion words.					
	Positive emotion words / #_comments	2.964	2.514	2.505	1.641	0.081
	Negative emotion words / #_comments	0.591	0.562	0.443	0.330	*
	H4: Fraudsters avoid using first-person pronouns, being reluctant to take ownership of their content.					
	Personal pronouns					
	First-person singular / #_comments	0.318	0.454	0.165	0.432	**
	First-person plural / #_comments	1.635	1.585	1.518	1.010	0.508
	Second-person / #_comments	1.757	1.562	1.674	1.010	0.610
	Third-person / #_comments	1.311	1.058	1.080	0.829	0.054
	H5: Fraudsters use fewer cognitive process words.					
	Exclusive words / #_comments	2.200	2.086	1.564	1.080	**
	Insight / #_comments	1.563	1.651	0.937	0.625	***
	Causal / #_comments	1.453	1.287	0.892	0.584	***
	H6: Fraudulent campaigns have a lower (text) readability.					
	Automated Readability Index	8.679	3.978	7.441	2.909	**
	Coleman-Liau Index	8.322	3.009	8.104	2.182	0.507
	Gunning-Fog Score	11.143	3.300	9.836	2.507	***
	Flesch-Kincaid Grade Level	7.472	3.001	6.170	2.357	***
	Flesch-Reading Ease score	71.626	16.701	77.508	14.728	**

APPENDIX B
DESCRIPTIVE ANALYSIS

Tables 6, and 7 provide descriptive analysis for the Campaign, Updates, and Comments sections, respectively. The tables are located at the end of the paper because of their large size.

REFERENCES

[1] *Kickstarter: Total Amount of Dollars Pledged 2022*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.statista.com/statistics/31

[2] C. Moores, "Kickstart my Lawsuit: Fraud and justice in rewards-based crowdfunding," *UC Philadelphia, PA, USA: Davis Law Rev.*, vol. 49, pp. 383–424, 2015.

[3] T. H. Ho, "Social purpose corporations: The next targets for greenwashing practices and crowdfunding scams," *Seattle J. Social Justice*, vol. 13, p. 935, Jan. 2015.

[4] J. Xu, D. Chen, and M. Chau, "Identifying features for detecting fraudulent loan requests on P2P platforms," *Proc. ISI*, vol. 38, pp. 79–84, Sep. 2016.

[5] M. Gorman and W. A. Sahlman, "What do venture capitalists do?" *J. Bus. Venturing*, vol. 4, no. 4, pp. 231–248, Jul. 1989.

[6] *Kobe Beef Jerky Kickstarter Fraud Nearly Scammed Backers Out of $120,000*. Accessed: Feb. 23, 2024. [Online]. Available: https://money.cnn.com/2013/06/17/technology/kickstarter-

[7] *Shitty Kickstarters*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.reddit.com/r/shittykickstarters/

[8] *Report Kickstarter Scams, IndieGogo Scams and Crowndfunding Creators & Scammers, Help Us Grow Our Knowledge Base and Keep Crowdfunding Accountable*. Accessed: Jul. 7, 2023. [Online]. Available: https://kickscammed.com/

[9] *Kickscammed: Crowdfunding Projects that Never Delivered*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.facebook.com/groups/1380253912299062/

[10] *Game Over: FTC Goes After Board Game Campaign Gone Wrong in First Crowdfunding Case*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.washingtonpost.com/news/the-switch/wp/2015

[11] *Crowdfunding Fraud: How Big is the Threat?*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.crowdfundinsider.com

[12] Q. Gao and M. Lin, "Economic value of texts: Evidence from online debt crowdfunding," in *Proc. Financial Innov., Online Lending Households Small Businesses*, Washington, DC, USA, 2016, pp. 1–4.

[13] M. Siering, J.-A. Koch, and A. V. Deokar, "Detecting fraudulent behavior on crowdfunding platforms: The role of linguistic and content-based cues in static and dynamic contexts," *J. Manage. Inf. Syst.*, vol. 33, no. 2, pp. 421–455, Apr. 2016.

[14] D. Cumming, L. Hornuf, M. Karami, and D. Schweizer, "Disentangling crowdfunding from fraudfunding," *J. Bus. Ethics*, vol. 182, no. 4, pp. 1103–1128, Feb. 2023.

[15] E. A. Smith and R. J. Senter, "Automated readability index," Aerosp. Med. Res. Lab., Univ. Cincinnati, Cincinnati, OH, USA, Tech. Rep., 1967.

[16] M. Coleman and T. L. Liau, "A computer readability formula designed for machine scoring," *J. Appl. Psychol.*, vol. 60, no. 2, pp. 283–284, Apr. 1975.

[17] R. Gunning, *The Technique of Clear Writing.*, 1st ed., New York, NY, USA: McGraw-Hill, 1952.

[18] R. Flesch, "A new readability yardstick," *J. Appl. Psychol.*, vol. 32, no. 3, pp. 221–233, 1948.

[19] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, Dept. Philosophy, Waikato Univ., Hamilton, New Zealand, 1999.

[20] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.

[21] C.-F. Tsai and Y.-C. Hsiao, "Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches," *Decis. Support Syst.*, vol. 50, no. 1, pp. 258–269, Dec. 2010.

[22] V. Etter, M. Grossglauser, and P. Thiran, "Launch hard or go home!: Predicting the success of kickstarter campaigns," in *Proc. 1st ACM Conf. Online Social Netw.*, Boston, MA, USA, Oct. 2013, pp. 177–182.

[23] M. W. Evers, "Main drivers of crowdfunding success," Erasmus Univ., Rotterdam, The Netherlands, Tech. Rep., 2012.

[24] J. Gera and H. Kaur, "Identifying significant features to improve crowd funded projects success," in *Innovations in Computer Science and Engineering*. Singapore: Springer, 2016, pp. 211–218.

[25] E. M. Gerber, J. S. Hui, and P. Y. Kuo, "Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms," in *Proc. CSCW*, Seattle, WA, USA, 2012, pp. 1–4.

[26] M. D. Greenberg, B. Pardo, K. Hariharan, and E. Gerber, "Crowdfunding support tools: Predicting success & failure," in *Proc. CHI EA*, Paris, France, Apr. 2013, pp. 1815–1820.

[27] T. Mitra and E. Gilbert, "The language that gets people to give: Phrases that predict success on kickstarter," in *Proc. CSCW*, Baltimore, MD, USA, Feb. 2014, pp. 49–61.

[28] E. Mollick, "The dynamics of crowdfunding: An exploratory study," *J. Bus. Venturing*, vol. 29, no. 1, pp. 1–16, Jan. 2014.

[29] F. Thies, M. Wessel, J. Rudolph, and A. Benlian, "Personality matters: How signaling personality traits can influence the adoption and diffusion of crowdfunding campaigns," in *Proc. ECIS*, Istanbul, Turkey, 2016, pp. 1–18.

[30] A. Xu, X. Yang, H. Rao, W.-T. Fu, S.-W. Huang, and B. P. Bailey, "Show me the money!: An analysis of project updates during crowdfunding campaigns," in *Proc. CHI*, New York, NY, USA, Apr. 2014, pp. 591–600.

[31] E. Mollick, "Delivery rates on Kickstarter," Social Sci. Res. Netw., Rochester, NY, USA, 2015. Accessed: Feb. 23, 2024. [Online]. Available: https://ssrn.com/abstract=2699251

[32] A. K. Agrawal, C. Catalini, and A. Goldfarb, "The geography of crowdfunding," Nat. Bur. Econ. Res., Cambridge, MA, USA, Tech. Rep., 2011.

[33] D. Zvilichovsky, Y. Inbar, and O. Barzilay, "Playing both sides of the market: Success and reciprocity on crowdfunding platforms," Social Sci. Res. Netw., Rochester, NY, USA, Tech. Rep., 2015. Accessed: Feb. 23, 2024. [Online]. Available: https://ssrn.com/abstract=2304101

[34] V. Rakesh, J. Choo, and C. K. Reddy, "Project recommendation using heterogeneous traits in crowdfunding," in *Proc. ICWSM*, Oxford, U.K., 2015, pp. 26–29.

[35] Antonio-M. Moreno-Moreno, C. Sanchís-Pedregosa, and E. Berenguer, "Success factors in peer-to-business (P2B) crowdlending: A predictive approach," *IEEE Access*, vol. 7, pp. 148586–148593, 2019.

[36] J. An, D. Quercia, and J. Crowcroft, "Recommending investors for crowdfunding projects," in *Proc. WWW*, Apr. 2014, pp. 261–270.

[37] M. Beier and K. Wagner, "Crowdfunding success of tourism projects-evidence from Switzerland," Social Sci. Res. Netw., Rochester, NY, USA, Tech. Rep., 2014. Accessed: Feb. 23, 2024. [Online]. Available: https://ssrn.com/abstract=2520925

[38] E. Mollick and R. Nanda, "Wisdom or madness? Comparing crowds with expert evaluation in funding the arts," *Manage. Sci.*, vol. 62, no. 6, pp. 1533–1553, Sep. 2015.

[39] T. Lynn, P. Rosati, B. Nair, and C. Mac an Bhaird, "An exploratory data analysis of the #Crowdfunding network on Twitter," *J. Open Innov., Technol., Market, Complex.*, vol. 6, no. 3, p. 80, Sep. 2020.

[40] K. Kim and S. Viswanathan, "The 'experts' in the crowd: The role of experienced investors in a crowdfunding market," Social Sci. Res. Netw., Rochester, NY, USA, Tech. Rep., 2018. Accessed: Feb. 23, 2024. [Online]. Available: https://ssrn.com/abstract=2258243

[41] W. Wang, L. He, Y. J. Wu, and M. Goh, "Signaling persuasion in crowdfunding entrepreneurial narratives: The subjectivity vs objectivity debate," *Comput. Hum. Behav.*, vol. 114, Jan. 2021, Art. no. 106576.

[42] J. K. Burgoon, D. B. Buller, K. Floyd, and J. Grandpre, "Deceptive realities: Sender, receiver, and observer perspectives in deceptive conversations," *Commun. Res.*, vol. 23, no. 6, pp. 724–748, Dec. 1996.

[43] J. K. Burgoon, J. P. Blair, T. Qin, and J. F. Nunamaker, "Detecting deception through linguistic analysis," in *Proc. ISI*, Tucson, AZ, USA, 2003, pp. 91–101.

[44] S. L. Humpherys, K. C. Moffitt, M. B. Burns, J. K. Burgoon, and W. F. Felix, "Identification of fraudulent financial statements using linguistic credibility analysis," *Decis. Support Syst.*, vol. 50, no. 3, pp. 585–594, Feb. 2011.

[45] P. S. Keila and D. B. Skillicorn, "Detecting unusual and deceptive communication in email," in *Proc. CASCON*, Toronto, ON, Canada, Jan. 2005, pp. 117–125.

[46] C. L. Toma and J. T. Hancock, "Reading between the lines: Linguistic cues to deception in online dating profiles," in *Proc. CSCW*, Savannah, GA, USA, 2010, pp. 5–8.

[47] J. Huang, G. Stringhini, and P. Yong, "Quit playing games with my heart: Understanding online dating scams," in *Proc. DIMVA*, Milan, Italy, 2015, pp. 216–236.

[48] L. Zhou and D. Zhang, "Following linguistic footprints: Automatic deception detection in online communication," *Commun. ACM*, vol. 51, no. 9, pp. 119–122, Sep. 2008.

[49] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language use: Our words, our selves," *Annu. Rev. Psychol.*, vol. 54, no. 1, pp. 547–577, Feb. 2003.

[50] L. Zhou, J. K. Burgoon, J. F. Nunamaker, and D. Twitchell, "Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications," *Group Decis. Negotiation*, vol. 13, no. 1, pp. 81–106, Jan. 2004.

[51] B. M. DePaulo, "Cues to deception," *Psychol. Bull.*, vol. 129, no. 1, p. 74, Jan. 2003.

[52] A. Vrij, *Detecting Lies and Deceit: The Psychology of Lying and Implications for Professional Practice*. Hoboken, NJ, USA: Willy, 2000.

[53] M. L. Newman, J. W. Pennebaker, D. S. Berry, and J. M. Richards, "Lying words: Predicting deception from linguistic styles," *Personality Social Psychol. Bull.*, vol. 29, no. 5, pp. 665–675, May 2003.

[54] W. Shafqat, S. Lee, S. Malik, and H.-C. Kim, "The language of deceivers: Linguistic features of crowdfunding scams," in *Proc. WWW*, Montreal, QC, Canada, 2016, pp. 99–100.

[55] S. Lee, W. Shafqat, and H.-C. Kim, "Backers beware: Characteristics and detection of fraudulent crowdfunding campaigns," *Sensors*, vol. 22, no. 19, p. 7677, Oct. 2022.

[56] *Kickstarter*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.kickstarter.com/

[57] *The Marketplace for Successfully Crowdfunded Projects*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.facebook.com/outgrowme/

[58] *Kickstarter's Top 50 Projects: When they Shipped-CNNMoney*. Accessed: Feb. 23, 2024. [Online]. Available: https://money.cnn.com/interactive/technology/kickstarter-

[59] *Selenium*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.selenium.dev/

[60] *NLTK : Natural Language Toolkit*. Accessed: Feb. 27, 2025. [Online]. Available: https://www.nltk.org/

[61] *The Stanford Natural Language Processing Group*. Accessed: Feb. 27, 2025. [Online]. Available: https://nlp.stanford.edu/software/tagger.html

[62] *Eye3 (Suspended) by Kellie Sigler Kickstarter*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.kickstarter.com/projects/eye3/eye3

[63] J. A. Koch, "Crowdfunding success factors: The characteristics of successfully funded projects on crowdfunding platforms," in *Proc. ECIS*, Muenster, Germany, 2015, pp. 1–15.

[64] W. Michael, T. Ferdinand, and B. Alexander, "A lie never lives to be old: The effects of fake social information on consumer decision-making in crowdfunding," in *Proc. ECIS*, Munster, Germany, 2015, pp. 1–15.

[65] A. Rauch and M. Frese, "Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success," *Eur. J. Work Organizational Psychol.*, vol. 16, no. 4, pp. 353–385, Dec. 2007.

[66] S. Shane and D. Cable, "Network ties, reputation, and the financing of new ventures," *Manage. Sci.*, vol. 48, no. 3, pp. 364–381, Mar. 2002.

[67] Y. Liu, "Word of mouth for movies: Its dynamics and impact on box office revenue," *J. Marketing*, vol. 70, no. 3, pp. 74–89, Jul. 2006.

[68] J. A. Chevalier and D. Mayzlin, "The effect of word of mouth on sales: Online book reviews," *J. Marketing Res.*, vol. 43, no. 3, pp. 345–354, Aug. 2006.

[69] M. Muckerman, B. M. Depaulo, and R. Rosenthal, "Verbal and nonverbal communication of deception," *Adv. Experim. soc. Psychol.*, vol. 14, pp. 1–59, Jan. 1981.

[70] *Welcome to LIWC-22*. Accessed: Feb. 27, 2025. [Online]. Available: https://www.liwc.app/

[71] *The Stanford Natural Language Processing Group*. Accessed: Jan. 3, 2023. [Online]. Available: https://nlp.stanford.edu/software/CRF-NER.html

[72] P. Ekman, *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*, 1st ed., New York, NY, USA: Norton, 2009.

[73] N. Liberman, Y. Trope, and E. Stephan, *Social Psychology*, 2nd ed., New York, NY, USA: The Guilford Press, 2007.

[74] M. Wiener and A. Mehrabian, *Language Within Language: Immediacy, a Channel in Verbal Communication*, 1st ed. New York, NY, USA: Irvington Pub, 1968.

[75] K. Moffitt and M. B. Burns, "What does that mean? Investigating obfuscation and readability cues as indicators of deception in fraudulent financial reports," in *Proc. AMCIS*, San Francisco, CA, USA, Jan. 2009, p. 399.

[76] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, Jul. 1997.

[77] *Weka 3-Data Mining with Open Source Machine Learning Software in Java*. Accessed: Feb. 23, 2024. [Online]. Available: https://www.cs.waikato.ac.nz/ml/weka/

[78] E. Rish, *Artificial Intelligence*, 2nd ed., New York, NY, USA: McGraw Hill, 1991.

[79] E. Rich and K. Knight, *Artificial Intelligence*. New York, NY, USA: McGraw-Hill, 1991.

[80] J. C. F. Winter, "Using the student's t-test with extremely small sample sizes," *Practical Assessment, Res., Eval.*, vol. 18, p. 10, Aug. 2013.

[81] G. D. Ruxton, "The unequal variance t-test is an underused alternative to student's t-test and the mann–whitney u test," *Behav. Ecol.*, vol. 17, no. 4, pp. 688–690, Jul. 2006.

[82] B. Kleinberg, M. Mozes, A. Arntz, and B. Verschuere, "Using named entities for computer-automated verbal deception detection," *J. Forensic Sci.*, vol. 63, no. 3, pp. 714–723, 2018.

[83] K. Papantoniou, P. Papadakos, T. Patkos, G. Flouris, I. Androutsopoulos, and D. Plexousakis, "Deception detection in text and its relation to the cultural dimension of individualism/collectivism," *Natural Lang. Eng.*, vol. 28, no. 5, pp. 545–606, Sep. 2022.

[84] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," 2011, *arXiv:1107.4557*.

[85] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Negative deceptive opinion spam," in *Proc. NAACL-HLT*, Atlanta, GA, USA, 2013, pp. 497–501.

[86] M. L. Knapp, R. P. Hart, and H. S. Dennis, "An exploration of deception as a communication construct," *Human Commun. Res.*, vol. 1, no. 1, pp. 15–29, Sep. 1974.

[87] J. Greenberg and E. Mollick, "Sole survivors: Solo ventures versus founding teams," Social Sci. Res. Netw., Rochester, NY, USA, Tech. Rep., 2018. Accessed: Feb. 23, 2024. [Online]. Available: https://ssrn.com/abstract=3107898

[88] N. Franke, M. Gruber, D. Harhoff, and J. Henkel, "Venture capitalists' evaluations of start–up teams: Trade–offs, knock–out criteria, and the impact of VC experience," *Entrepreneurship Theory Pract.*, vol. 32, no. 3, pp. 459–483, May 2008.

[89] S. M. Ho and J. T. Hancock, "Context in a bottle: Language-action cues in spontaneous computer-mediated deception," *Comput. Hum. Behav.*, vol. 91, pp. 33–41, Feb. 2019.

[90] M. L. Scott and S.-I. Lee, "A unified approach to interpreting model predictions," 2017, *arXiv:1705.07874*.

SEUNGHUN LEE received the B.S. and M.S. degrees in computer science and engineering from Sangmyung University, Cheonan, South Korea, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree in software. His research interests include data mining, internet financial platforms, and natural language processing.

HEEMIN PARK received the B.S. and M.S. degrees in computer science from Sogang University, South Korea, in 1993 and 1995, respectively, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 2006. He was with Yonsei University, Sookmyung Women's University, and Samsung Electronics, South Korea. He is currently a Professor with the Department of Computer Software Engineering, Sangmyung University, Cheonan, South Korea. His research interests include web-based information systems, networked and embedded computing systems, cyber-physical systems, multimedia applications, and entertainment computing.

HYUN-CHUL KIM (Member, IEEE) received the Ph.D. degree from KAIST in 2005. He is a Professor with the Department of Software, Sangmyung University. He was with CAIDA, UCSD, and Seoul National University. His research interests focus on data-driven analysis of the structure and collective dynamics of large-scale networked systems, with an emphasis on enhancing their scalablity, performance, security, safety, and privacy.