Small Bus Econ
https://doi.org/10.1007/s11187-024-00991-3

RESEARCH ARTICLE

Signaling and perceiving on equity crowdfunding decisions — a machine learning approach

Jinjuan Yang · Jiayuan Xin · Yan Zeng · Pei Jose Liu

Accepted: 4 December 2024

Abstract This study explores how signaling and perceiving jointly influence crowd investors' decision-making. We utilize five machine learning models to assess the predictive power of various information types on crowdfunding success. Our findings indicate that investors prioritize well-structured quantitative data over complex qualitative content. Processing quantitative information is also found to be less cognitively taxing than extracting useful information from qualitative text and images. Entrepreneurs' signaling and investors' processing jointly reduce information asymmetry in crowdfunding, highlighting the critical yet often-overlooked role of investors' information processing. Additionally, we test the policy effect of the '*2016 Interim Measures on Online Lending*' on crowdfunding success by comparing the predictive accuracy of information during the thriving and constraining periods of crowdfunding development in China. Our results have significant implications for policymakers that crowdfunding fosters economic growth by connecting entrepreneurs and investors and should not be halted due to risks, especially during periods of financial constraints.

J. Yang
School of Business and Management, Shanghai International Studies University, Shanghai, China

J. Xin · P. J. Liu (✉)
Newcastle University Business School, Newcastle University, Newcastle Upon Tyne, UK
e-mail: pei.liu@newcastle.ac.uk

Y. Zeng
Faculty of Business and Law, University of the West of England, Bristol, UK

Plain English Summary This study examines how the information shared by entrepreneurs and processed by investors impacts crowdfunding success. Using five machine learning models, we find that investors prioritize clear, well-structured quantitative data, as it is easier to process than text or images. Presenting information in a way that reduces cognitive effort helps bridge the gap between entrepreneurs and investors. We also analyze the impact of China's *'2016 Interim Measures on Online Lending*,' finding that while regulations stabilize crowdfunding, shutting platforms down harms economic growth, particularly during financial instability. The principal implication of this study is that policymakers should regulate crowdfunding to address risks without stifling its potential. Crowdfunding platforms provide essential funding opportunities for small business entrepreneurs, especially in constrained financial environments.

Keywords Equity crowdfunding · Information asymmetry · Dual-system · Cognitive processing · Machine learning

JEL Classification G11 · G21 · G32 · D82

Published online: 18 January 2025

1 Introduction

Startups often struggle to secure financing from traditional banks due to their early-stage and uncertain nature. For decades, Silicon Valley Bank (SVB) played a pivotal role in bridging this funding gap, supporting nearly half of all the U.S. venture-backed firms and significantly contributing to the growth of the technology industry both domestically and internationally. The sudden collapse of SVB has left a considerable void in startup financing, raising concerns among founders and investors about the survival of high-potential startups in an already constrained funding environment. In this context, equity crowdfunding,[1] has emerged as an alternative source of capital, enabling unlisted startups to raise funds from a broader base of investors (Cumming et al., 2019b; De Crescenzo et al., 2020; Wang et al., 2019). However, a critical challenge in crowdfunding investment lies in the high-noise operating environment, where entrepreneurs often present a bundle of diverse signals to attract potential backers (Courtney et al., 2017; Plummer et al., 2016). This practice can overwhelm investors, particularly those with less experience compared to venture capitalists and angel investors (Colombo et al., 2019; Steigenberger & Wilhelm, 2018; Vismara, 2019).

Our research addresses the question: How do signaling strategies and investor perception interact to influence crowdfunding success in such a high-noise environment? This study adopts a unique approach by integrating cognitive signaling theory with machine learning models to explore both the signaling behavior of entrepreneurs and the cognitive processing of crowd investors. While signaling theory has traditionally focused on how entrepreneurs reduce information asymmetry by signaling project quality, it has largely neglected the perspective of signal receivers—crowd investors—who are tasked with interpreting and making sense of multiple, often conflicting signals (Drover et al., 2018). Our study fills this gap by examining how different types of information are processed by investors and how this affects their investment decisions.

Innovatively, we employ machine learning models to simulate investors' cognitive processing modes. Previous research has shown that human decision-making can be emulated using computational models that replicate heuristic and systematic processing. We propose that well-structured quantitative information can be effectively processed through a decision tree (DT) approach, which relies on domain knowledge and requires less cognitive effort (Sugumaran & Ramachandran, 2007; Tran et al., 2009). In contrast, complex and ambiguous qualitative information, such as narrative texts and images, demands more sophisticated cognitive abilities, which can be modeled using a backpropagation neural network (BP-NN) that mimics the human brain's neural system (Bayoudh et al., 2022; Brown et al., 1993; Curram & Mingers, 1994). By manipulating the cognitive efforts of crowdfunding investors using DT and BP-NN models, we explore how investors process and respond to various signals, contributing to a more nuanced understanding of decision-making in crowdfunding.

Our methodological framework is inspired by the transformative opportunities that machine learning brings to digital finance. By integrating advanced computational techniques, sophisticated algorithms, and extensive data resources, ML delivers highly accurate, data-driven insights that significantly enhance and refine investment decision-making processes. For instance, Meoli and Vismara (2022) highlight ML's predictive accuracy in forecasting ICO success by analyzing structured data like financial metrics and unstructured inputs such as social media sentiment, uncovering trends that reduce uncertainty in the volatile cryptocurrency market. Additionally, ML techniques demonstrate the potential to emulate key aspects of human decision-making, encompassing both subconscious and conscious processes. Subconscious decision-making, characterized by swift and intuitive evaluations, parallels ML's ability to rapidly analyze extensive datasets and detect subtle

[1] Equity crowdfunding refers to a financing mode based on internet channels in which a company transfers a certain percentage of its shares to ordinary investors, who can obtain future profits by investing in the company. Other than equity crowdfunding, there are three other categories for crowdfunding: in reward crowdfunding, crowd investors donate to a project or business with the expectation of receiving a non-financial reward in return, such as goods or services at a later stage; in lending crowdfunding, crowd investors receive interest on their investment; in donation crowdfunding, crowd investors receive no reward in exchange for their contributions but rather gain the personal satisfaction of supporting a project that they consider meaningful and worthy (Cappa et al., 2020).

patterns (LeCun et al., 2015; Mullainathan & Spiess, 2017). For instance, ML can process unstructured data, such as social media sentiment or behavioral trends, to generate insights akin to human intuition. Similarly, conscious reasoning, involving deliberate and analytical thought, is mirrored by advanced ML algorithms designed for logical problem-solving and optimization (LeCun et al., 2015; Tenenbaum et al., 2011).

Another key aspect of our research is the impact of regulatory reforms on crowdfunding, with a particular focus on the Chinese market, which offers a unique setting to examine the effects of policy changes in an emerging economy. In many developing regions, rapid growth of crowdfunding platforms, coupled with weak regulatory frameworks, has led to significant risks for investors, entrepreneurs, and regulators, particularly around illegal fundraising (Huang et al., 2018). China, with its early light-touch regulation, saw a surge in fraudulent activities, culminating in high-profile cases like the Ezubao scandal, one of China's largest peer-to-peer lending platform,[2] which defrauded investors of RMB 50 billion ($8 billion). This crisis prompted the introduction of the *"2016 Interim Measures on Online Lending"* (hereafter referred to as the *Interim Measures*), which imposed stricter information disclosure and due diligence requirements, transforming the regulatory landscape. Studying China's experience provides valuable insights into how regulatory interventions can enhance transparency and credibility in high-risk markets, offering lessons that can be generalized to other emerging economies where digital finance outpaces regulatory oversight.

We collect a unique dataset comprising 144 projects from August 2014 to January 2019 on Dreammove, a prominent equity crowdfunding platform in China. Our study reveals that crowd investors exhibit a clear preference for well-structured and easily interpretable quantitative information over more complex qualitative signals. Moreover, we find that the cognitive effort required to process quantitative numerical data is lower than that needed to interpret qualitative texts and images, suggesting that the complexity of the information significantly influences the level of analytical effort required. Effective signaling by entrepreneurs and efficient information processing by investors are both critical in reducing information asymmetry in crowdfunding. We further find that the introduction of the 2016 *Interim Measures* has mitigated the impact of excessive noise in qualitative signals that could distract crowd investors. This regulatory intervention has created a more structured environment, allowing investors to focus on explicit, well-organized quantitative information, thereby supporting more informed decision-making.

Our research makes several key contributions. First, we assess the impact of synergy and conflict within a signal portfolio on decision-making in equity crowdfunding. While signals interact to capture investor attention, this interaction remains underexplored (Bapna, 2019; Kleinert & Vismara, 2023). Our findings reveal that conflicting signals often reduce decision accuracy through a "cancel out" effect, emphasizing the need for clarity and consistency in signaling. Entrepreneurs and platforms should consider separating signals for independent evaluation to improve decision-making (Steigenberger & Wilhelm, 2018). Second, we extend signaling theory by introducing two cognitive dimensions: heuristic vs. systematic processing and holistic vs. segmental processing. Heuristic processing is fast and intuitive, ideal for managing information overload (Gigerenzer & Gaissmaier, 2011), while systematic processing requires deeper, more deliberate analysis, essential for complex signals (Chaiken & Ledgerwood, 2012). Holistic processing integrates signals into a unified assessment, whereas segmental processing breaks them down into categories (Laureiro-Martínez & Brusoni, 2018). Our framework highlights how the balance between these dimensions influences decision-making effectiveness, which varies depending on the context and type of information (Hoegen et al., 2018; Vismara, 2018). Third, our use of machine learning to model cognitive processes in signal interpretation opens new methodological possibilities. Unlike traditional research that relies on experiments, which may be limited by participant differences and contextual variability (Klein et al., 2018; Schneider & Harknett, 2019), machine learning provides a scalable and data-driven framework for analyzing how people process information and make decisions. Decision trees (DT) and backpropagation neural

[2] The peer-to-peer (P2P) online lending is regarded as loan-based crowdfunding, different from investment-based crowdfunding (Huang, 2018).

networks (BP-NN) replicate heuristic and systematic processing, respectively, offering a nuanced approach to understanding investor behavior in complex decision-making environments. Finally, we explore the impact of regulatory reforms on signal effectiveness in equity crowdfunding, offering a novel exploration of regulatory effects in transitional economies. The 2016 *Interim Measures*, while primarily targeting peer-to-peer lending, have improved signal quality in equity crowdfunding by enforcing due diligence, disclosures, and risk management (Ding et al., 2021; Wang et al., 2019). Our findings indicate that these regulations enhanced signal credibility and investor decision-making, underscoring the critical role of regulatory oversight in promoting market stability and growth (Huang, 2018).

This paper is organized as follows. Section 2 discusses the theoretical framework. Section 3 reviews the related literature and develops our hypothesis. Section 4 describes the methodology. Section 5 introduces the sample and measurements of input and output variables. Section 6 presents and analyzes the main results. Section 7 discusses various robustness tests, and Section 8 concludes.

2 Theoretical framework

2.1 Signaling theory in entrepreneurial finance

Signaling theory addresses information asymmetry in uncertain markets by enabling high-quality ventures to distinguish themselves through credible signals. These signals create a separating equilibrium, where only ventures with superior attributes can afford to emit costly signals, thereby reducing information asymmetry and aiding stakeholders in making more informed investment decisions (Bergh et al., 2014; Spence, 1973). This is particularly relevant in entrepreneurial finance, where new ventures often struggle to convey their true quality to potential investors due to limited track records and visibility (Plummer et al., 2016; Reuer et al., 2012). Early-stage ventures rely on various signaling mechanisms, such as securing patents, endorsements from reputable partners, the acquisition of patents, or the founders' previous successes, to effectively communicate their value, reduce perceived risks, and attract the necessary resources for growth (Ahlers et al., 2015; Busenitz et al., 2005; Colombo et al., 2019; Piva & Rossi-Lamastra, 2018). These signals act as credible indicators of a venture's quality, significantly increasing the likelihood of securing investment and enhancing long-term success prospects.

The literature in entrepreneurial finance categorizes signals into costly and costless, entrepreneurial and project-related, and value versus commitment signals. Costly signals, such as securing patents, retaining equity, or forming strategic partnerships, are more credible due to their substantial resource requirements, making them challenging for lower-quality ventures to replicate (Ahlers et al., 2015; Vismara, 2016). Costless signals, though easier to produce and generally less credible, can still influence perceptions by reflecting the entrepreneur's vision and strategic intent (Di Pietro et al., 2023; Kleinert & Volkmann, 2019).

Entrepreneurial signals, like the founder's track record and industry reputation, are crucial for assessing leadership competence and reliability, which are important for investors evaluating the ability to navigate challenges (Busenitz et al., 2005). Project-related signals, such as product certifications, customer endorsements, or working prototypes, provide tangible evidence of a venture's operational viability, addressing concerns about technical feasibility and market readiness (Butticè et al., 2022).

Value signals, such as detailed financial projections, market analysis, and business plans, highlight a venture's intrinsic worth through detailed financial projections, market analysis, and business plans that showcase potential for profitability and growth (Cumming et al., 2015). Commitment signals, such as equity retention by founders or the formation of long-term strategic partnerships, demonstrate the founders' dedication to the venture's success, aligning their interests with those of potential investors (Vismara, 2016).

Despite these useful categories, they do not fully account for the dynamic interplay between signals or the varied interpretations by different stakeholders. The potential synergy, complementarity, or conflict among signals can significantly influence their overall impact, yet this aspect remains underexplored (Bapna, 2019; Busenitz et al., 2005; Piva & Rossi-Lamastra, 2018). Additionally, the cognitive processing of signal receivers plays a crucial role in determining how signals are interpreted and how effectively information is transmitted. Understanding these differences is essential for tailoring signals to meet the specific expectations and priorities of various investors.

2.2 A cognitive perspective on signaling theory

While selective attention has been acknowledged as a key mechanism for managing information overload, the deeper cognitive processes involved in interpreting and integrating these signals into final investment decisions remain underexplored (Drover et al., 2018). Traditional signaling theory assumes homogeneous interpretations of signals by investors, overlooking the potential for diverse perceptions. Drover et al. (2018) addresses this gap by introducing cognitive signaling theory, which incorporates a dual-system model of cognitive processing. This model differentiates between heuristic processing, which is low-effort, automatic, and intuitive, and systematic processing, which is more deliberate, analytical, and resource-intensive (Evans, 2006, 2008, 2010).

Dane and Pratt (2007) suggest that heuristic and systematic processing often operate together during decision-making. Individuals may initially rely on heuristic processing but engage in systematic processing when detailed analysis is needed. Griffin et al. (2004), Chen et al. (1999), and Ferran and Watts (2008) support this, noting that decision-makers switch to systematic processing when heuristic processing does not yield sufficient confidence. The choice of processing method depends on various factors, including the investor's capabilities, motivations, and the nature of the signals. Todd and Gigerenzer (2012) argue that individuals with limited knowledge or experience are more likely to rely on heuristics for simplified decision-making. Similarly, Edelman et al. (2021) find that business angels tend to use heuristics when time constraints demand swift decisions. Signal complexity also plays a critical role. Tversky and Kahneman (1974) suggest that complex or uncertain signals encourage heuristic processing as a way to manage information overload, though this may come at the cost of accuracy. Conversely, Vergne et al. (2018) highlight that true complexity pushes investors toward more systematic, analytical processing.

2.2.1 *Heuristic vs. systematic processing*

In crowdfunding, ventures often present multiple signals simultaneously to attract a broad range of investors (Plummer et al., 2016; Steigenberger & Wilhelm, 2018). Investors typically process this bundled information holistically, using heuristic approaches for quick decision-making (Tversky & Kahneman, 1974). This strategy allows investors to make rapid judgments by simplifying complex information. However, when signals are diverse and complex, they may overwhelm the cognitive capacity for fast, intuitive judgments, requiring more systematic, deliberate processing (Laureiro-Martínez & Brusoni, 2018).

The question of whether heuristic or systematic processing leads to superior decisions in crowdfunding remains underexplored (Hoegen et al., 2018; Vismara, 2018). Although heuristics can lead to errors, as suggested by Tversky and Kahneman (1974), the concept of "ecological rationality" posits that heuristics can yield better decisions in uncertain environments (Todd & Gigerenzer, 2012). Heuristics are especially effective under time pressure and information overload, which are common in crowdfunding (Hoegen et al., 2018; Mollick, 2014). Systematic processing, on the other hand, involves detailed analysis of project signals, leading to more stable investment decisions (Allison et al., 2017; Bi et al., 2017). While it can provide a more comprehensive evaluation, it may also lead to information overload or decision paralysis, common in crowdfunding platforms (Eppler & Mengis, 2004; Klein, 2015). Drover et al. (2018) suggests that ambiguous signals prompt systematic processing, while clearer signals may encourage heuristic processing, allowing for quicker decisions.

Thus, the effectiveness of heuristic versus systematic processing in crowdfunding is context-dependent, shaped by signal clarity, complexity, and the decision-maker's expertise (Drover et al., 2018). Heuristic processing is fast and efficient, but may overlook critical details, whereas systematic processing is more thorough but time-consuming, potentially leading to decision delays. The choice between these approaches requires a balance between speed and accuracy, depending on the specific context and investor expertise (Chaiken & Ledgerwood, 2012; Gigerenzer & Gaissmaier, 2011; Hoegen et al., 2018; Vismara, 2018).

2.2.2 *Segmental vs. holistic processing*

Beyond cognitive depth (heuristic vs. systematic), we must also consider the scope of information processing (segmental vs. holistic). These two dimensions together help identify the cognitive strategies best suited to navigating the information landscape of equity crowdfunding. Segmental processing focuses on analyzing individual signals in isolation, which can help allocate cognitive resources more efficiently, improve venture quality assessments, and reduce biases (Kahneman & Tversky, 2013; Payne et al., 1988). This approach is effective for complex decisions, allowing investors to examine both quantitative and qualitative factors separately for more balanced judgments (Franke et al., 2008; Slovic, 1972).

On the other hand, holistic processing—viewing all signals together—can lead to faster, automatic judgments and align with bounded rationality and cognitive shortcuts in complex decision-making (Gigerenzer & Gaissmaier, 2011; Simon, 1990). In high-noise environments like crowdfunding, holistic processing helps investors rapidly integrate diverse signals, reducing information asymmetry, and boosting decision efficiency (Kahneman, 2011; Steigenberger & Wilhelm, 2018). However, Laureiro-Martínez and Brusoni (2018) argue that the complexity of these signals can overwhelm investors' capacity for fast, intuitive judgments, making systematic analysis necessary for more accurate evaluations.

Although both segmental and holistic processing have their merits, in equity crowdfunding, the balance often leans toward holistic-heuristic processing. Given the high volume of information and limited time for in-depth analysis, holistic processing allows investors to synthesize diverse signals more efficiently and make quicker decisions (Correia et al., 2024). This method also helps investors assess more opportunities and build diversified portfolios, ultimately enhancing their decision-making effectiveness (Barber & Odean, 2008).

3 Literature review and hypothesis development

3.1 Signaling in crowdfunding

Crowdfunding, particularly equity crowdfunding (ECF), has become a vital component of entrepreneurial finance, enabling ventures to raise capital from small-scale private investors through digital platforms (Butticè & Vismara, 2022). Characterized by high information asymmetry and uncertainty, ECF presents a challenging environment for both ventures and investors (Ahlers et al., 2015; Vismara, 2016). In this high-noise setting, early-stage ventures must strategically deploy signal portfolios to capture investor attention and convey quality (Courtney et al., 2017; Plummer et al., 2016). However, investors, overwhelmed by information, often rely on selective attention mechanisms, focusing on signals they perceive as most relevant and credible (Butticè et al., 2022).

3.1.1 *Quantitative signals*

Quantitative signals, such as financial projections, accounting ratios, and market analysis, because they provide concrete, measurable data directly tied to a crowdfunding project's potential profitability and growth. These signals are essential as they offer a factual basis for evaluating a project's financial health, allowing investors to make informed decisions with greater confidence. By reducing uncertainty, these quantitative signals play a crucial role in attracting investor attention, as they enable objective analysis and comparison across different projects (Ahlers et al., 2015; Cohen et al., 2020).

In addition, equity retention and strategic partnerships, often viewed as high-credibility signals, also draw significant attention from investors. These signals are particularly powerful because they provide tangible evidence of a project's reliability and commitment to long-term success. For instance, the percentage of equity retained by founders or the terms of strategic partnerships can be quantified, offering clear, measurable indicators of the venture's stability and the founders' vested interest in its success (Vismara, 2016). These quantifiable elements make high-credibility signals more persuasive and easier for investors to evaluate, further enhancing their appeal.

A founder's track record and industry reputation, considered contextually relevant signals, are equally crucial for investors who value leadership quality and strategic alignment (Busenitz et al., 2005; Chen et al., 2009). Presented in quantifiable formats, such as the number of successful ventures previously led by the founder, years of experience in the industry, or the market share held by the venture, these signals provide concrete evidence of the leadership's competence and execution potential, which is particularly important for investors seeking to assess the team's capability to navigate challenges and drive growth (Bapna, 2019). In contrast to qualitative narratives, which are open to interpretation, quantitative data offers a clear, objective basis for evaluating a venture, thereby enhancing its appeal to investors.

3.1.2 Qualitative signals

Qualitative signals, including narrative texts and visual images, present a more complex picture. Narratives, such as descriptions of a firm's business activities or the founder's vision, often convey intricate information that requires sophisticated interpretation (Bapna, 2019; Monin et al., 2013). These narratives can be embellished or overly optimistic, adding to their ambiguity (Asay et al., 2018). This complexity may require investors to engage in more systematic cognitive processing (Franzoni & Tenca, 2023; Laureiro-Martínez & Brusoni, 2018). Meanwhile, visual content, such as leadership images, can engage investors emotionally by leveraging the "beauty premium," where physical attractiveness influences perceptions of trustworthiness, reliability, and authenticity (Colombo et al., 2022; Cook & Mobbs, 2022). These qualitative signals, while less concrete, can complement quantitative elements by drawing attention to a venture's strengths and engaging investors on an emotional and strategic level (Steigenberger & Wilhelm, 2018). When combined with quantitative data, qualitative signals can frame the investment proposition within a compelling narrative, making the project more attractive (Bafera & Kleinert, 2023; Bapna, 2019). Given the strengths and weaknesses of both quantitative and qualitative signals, we propose the following hypothesis:

> Hypothesis 1a: Investors tend to prioritize quantitative signals over qualitative ones, making quantitative signals more effective in predicting campaign success.

3.1.3 Signal bundle

However, the interaction between qualitative and quantitative signals can sometimes create tension. For example, overly optimistic narratives that lack solid quantitative backing may generate skepticism among data-driven investors. As Kleinert and Volkmann (2019) note, qualitative signals gain credibility when supported by substantial quantitative evidence. Maintaining a balance between these two types of signals is essential; overreliance on either may fail to comprehensively capture investor interest or inspire confidence. The success of a signal portfolio depends on how well the qualitative narrative aligns with the quantitative data, addressing investors' diverse concerns and motivations.

Our study is the first to explore whether the synergy between qualitative and quantitative signals enhances crowdfunding success or whether a disconnect between these signals weakens their impact. Entrepreneurs who craft signal portfolios where qualitative and quantitative signals reinforce each other present a cohesive, persuasive case that resonates with investors, reducing information asymmetry and boosting confidence. Conversely, a mismatch between these signals can undermine credibility and reduce the overall effectiveness of the portfolio.

> Hypothesis 1b: The predictive accuracy of a signal portfolio depends on the interplay between quantitative and qualitative signals, with synergy enhancing campaign success and conflicts diminishing its effectiveness.

3.2 Cognitive processing in crowdfunding

Building on our previous signal classification, we examine two signal categories: quantitative and qualitative (narrative texts and visual images). We argue that investors apply different cognitive processing depending on the context. This tailored approach enables investors to more effectively assess each signal's contribution and risk, leading to more accurate predictions of campaign success.

3.2.1 Quantitative signals

We posit that measurable signals, particularly quantitative data, offer greater clarity due to their grounding in concrete numbers and alignment with widely understood concepts such as credibility and leadership quality (Bapna, 2019; Vismara, 2016). These signals often conform to established heuristics, facilitating intuitive, rule-based decision-making (Gigerenzer & Gaissmaier, 2011). By enabling straightforward comparisons against benchmarks, quantitative data reduces cognitive load and minimizes subjective interpretation, making it a reliable tool for decision-making (Shah & Oppenheimer, 2008).

However, when investors lack the expertise to interpret technical data, such as financial projections, the objectivity of these numbers can be compromised, leading to misinterpretation or oversimplification (Drover et al., 2018; Kahneman, 2011). In our study, the quantitative data primarily focus on credibility, commitment, and leadership quality, which are aligned with common-sense principles rather than requiring specialized knowledge. As a result, these signals are more likely to trigger intuitive responses with less analytical effort, improving campaign success predictions.

> Hypothesis 2a: Quantitative (numerical) signals trigger intuitive responses via mental shortcuts, leading to more accurate predictions of campaign success than systematic analysis, provided investors have a sufficient knowledge base.

3.2.2 *Qualitative signals*

The analysis of narratives requires investors to learn and reflect in order to absorb multiple features, understand the sequence of events, and grasp the underlying relationships and meanings. This process involves reflecting on both chronological and non-chronological aspects, as well as dealing with ambiguity (Franzoni & Tenca, 2023; Mishler, 1995). Such in-depth analysis can lead to a deeper understanding of the investment opportunity and potentially more accurate predictions of campaign success.

However, excessive complexity in narratives can sometimes overwhelm investors, leading to an "underproduction" of interpretations (Hirshleifer & Teoh, 2003; Kahneman, 2011). In crowdfunding, this concern is mitigated as entrepreneurs often avoid overly complex narratives due to platform constraints and the need to engage a diverse investor base with simpler, more accessible content (Courtney et al., 2017; Kunz et al., 2017; Mollick, 2014). Thus, while narratives demand more cognitive effort, they can still contribute to accurate decision-making when appropriately designed.

Processing visual signals, particularly images, typically engages more intuitive and automatic cognitive mechanisms, tapping into deeply ingrained neural frameworks related to facial recognition and visual perception (Kanwisher & Yovel, 2006; Tsao & Livingstone, 2008). This allows the brain to process visual information rapidly, often bypassing more deliberate cognitive processes (Evans, 2008; Kahneman, 2011). However, the beauty premium is more complex than straightforward quantitative data due to its reliance on cognitive biases (Tversky & Kahneman, 1974), the variability of visual cues (Laureiro-Martínez & Brusoni, 2018), and contextual influences (Steigenberger & Wilhelm, 2018), making it harder to define and standardize.

The complexity of interpreting visual cues increases when investors focus on microfacial expressions, which can be subtle and ambiguous. These expressions may reveal underlying emotions or intentions that are not immediately apparent, prompting more detailed and analytical scrutiny (Yan et al., 2013; Zhi et al., 2021). Even when investors are aware of the beauty premium and try to look beyond superficial attributes, interpreting facial cues still demands higher cognitive effort to assess the founder's true emotions or intentions. Thus, while visual content can be processed quickly and intuitively, it also has the potential to require deeper analysis when investors seek to uncover less obvious signals.

> Hypothesis 2b: Qualitative signals, such as narrative texts and visual images, with their complexity and ambiguity, are more accurately processed through systematic analysis, leading to better predictions of campaign success.

3.2.3 *Signal bundle*

In high-noise crowdfunding environments, investors must often process a diverse array of signals under time constraints and limited cognitive resources (Correia et al., 2024). Holistic processing allows investors to synthesize multiple signals into a coherent impression, enhancing decision-making efficiency (Gigerenzer & Gaissmaier, 2011). This approach reduces information asymmetry by enabling investors to quickly integrate signals, improving campaign success predictions (Ahlers et al., 2015). By lowering cognitive load, holistic-heuristic processing allows investors to evaluate more opportunities and build more diversified portfolios, improving overall investment outcomes (Barber & Odean, 2008).

However, holistic processing can increase cognitive load when integrating diverse signals like quantitative and qualitative data (Chandler & Sweller, 1991; Paas et al., 2003), as task-switching between different information types is mentally taxing (Braver et al., 2003; Monsell, 2003). This challenges the notion that limited attention always favors holistic processing. In contrast, segmental processing proves more effective in complex decisions by allowing focused analysis of each signal type (Shanteau, 1992). It optimizes cognitive resources, reduces biases (Kahneman & Tversky, 2013), and leads to more accurate investment decisions through thorough evaluation of both quantitative and qualitative data (Franke et al., 2008; Slovic, 1972).

> Hypothesis 2c: In high-noise crowdfunding environments, holistic processing of signal bundles triggers intuitive, automatic responses, leading to quicker decision-making and more efficient predictions of campaign success compared to systematic processing.

3.3 Regulatory and environment uncertainties

Regulation plays a critical role in shaping the decision-making environment in equity crowdfunding by reducing information asymmetry and enhancing signal credibility. As the signaling environment rapidly evolves due to digitalization and external shocks, regulations help stabilize the investment landscape, enabling investors to adapt their strategies more effectively (Huang et al., 2022). Well-designed regulatory frameworks mitigate the high-noise environment typical of crowdfunding platforms (Mahmood et al., 2019) by standardizing information disclosure and enhancing transparency (Ahlers et al., 2015; Vismara, 2018). This standardization facilitates more efficient comparisons of investment opportunities (Rossi et al., 2021).

The introduction of the "*Interim Measures for the Administration of the Business Activities of Online Lending Information Intermediary Institutions*" in 2016 by the Chinese government illustrates how regulation can significantly impact equity crowdfunding. Although primarily targeting online lending, these measures also influenced equity crowdfunding by establishing a comprehensive regulatory framework. The *Interim Measures* addressed the rapid growth and risks in the online finance sector by setting guidelines on platform operations, permissible activities, and responsibilities (Ding et al., 2021; Huang, 2018). For equity crowdfunding, these regulations enhanced transparency and investor protection by mandating due diligence, financial disclosure, and robust risk management practices, thereby indirectly shaping the sector (Wang et al., 2019). This regulatory framework improved signal credibility, aligning with signaling theory by reducing information asymmetry between entrepreneurs and investors (Ahlers et al., 2015; Vismara, 2018).

In China's emerging equity crowdfunding market, these regulatory improvements are crucial in addressing challenges similar to those seen in the P2P lending crisis, such as poor project quality and inadequate oversight. The regulations help reduce market noise by enforcing stricter information disclosure, providing more reliable and transparent data for investors. This enhances the predictive accuracy of crowdfunding success, allowing for better assessment of project attributes, financial metrics, and entrepreneurial expertise. The outcome is more precise risk assessments, increased investor confidence, and a more efficient and effective equity crowdfunding environment in China.

> Hypothesis 3: Following the implementation of the Interim Measures, the reduction in environmental noise led to improved information quality, thereby enhancing the effectiveness of predicting campaign success compared to the pre-regulation period.

4 Methodology

Drover et al. (2018) call for empirical investigations into the dual-system processing model within signaling theory. While traditional cognitive science experiments, often involving undergraduates, have tested heuristic-systematic models (Bago et al., 2020; Van Bavel & Pereira, 2018), such controlled settings may not fully represent the diverse crowd in crowdfunding platforms (Anglin et al., 2018; Danilov & Sliwka, 2017). This demographic diversity makes it challenging to manipulate information processing using designed stimuli, as individuals from varied backgrounds may not conform to controlled experimental conditions. Moreover, the dual-system approach to human cognition is more complex than a simple binary of heuristic or systematic thinking; these processes often operate simultaneously and interactively (Alter et al., 2007; Dane & Pratt, 2007). Therefore, laboratory experiments may oversimplify the complex cognitive mechanisms in real-world decision-making contexts like crowdfunding.

4.1 Incorporating machine learning in cognitive modeling

Inspired by literature on organizational decision-making, which leverages machine learning (ML) models to simulate human cognition and decision-making (Kahneman, 2011; Phan et al., 2017), we explore how ML can replicate the cognitive processes involved in signaling theory. Cognitive modeling seeks to understand human thought processes through introspection (observing our own thoughts) and empirical observation (psychological experiments), allowing researchers to develop computational models that mimic human reasoning (Evans, 2008; Gigerenzer & Gaissmaier, 2011; Simon, 1990). By comparing the outcomes of these models with human behavior, we gain insights into cognitive processes and decision-making strategies (Eriksson et al., 2020; Haefner et al., 2021; Todd & Gigerenzer, 2012). This approach enhances our understanding of investor behavior in crowdfunding, extending the applicability of cognitive research from controlled experiments to more dynamic, real-world settings (Drover et al., 2018; Laureiro-Martínez & Brusoni, 2018).

4.2 Comparing decision trees and neural networks in cognitive simulation

To emulate dual-system cognitive processing in signaling theory, we employ decision trees (DT) and neural network (NN) models. Decision trees are well-suited for heuristic processing as they rely on quick, rule-based decisions using a straightforward if–then logic, which mirrors how humans make decisions based on specific criteria (Sugumaran & Ramachandran, 2007; Tran et al., 2009). Decision trees are hierarchical models that represent data through a series of binary splits based on feature values, with each node representing a decision rule and each leaf representing an outcome (see Fig. 1). They are intuitive and align with human decision-making processes, making them suitable for interpreting well-structured data (Gigerenzer & Gaissmaier, 2011; Hagras, 2018). However, DTs are less effective for complex, high-dimensional data due to their inability to capture intricate relationships and their rigidity in structure once built (Breiman et al., 2017). We employ DTs to model quick, rule-based processing in evaluating straightforward quantitative signals.

In contrast, neural networks, particularly back-propagation neural networks (BP-NNs), replicate systematic processing by modeling intricate, non-linear relationships akin to the human brain's neural

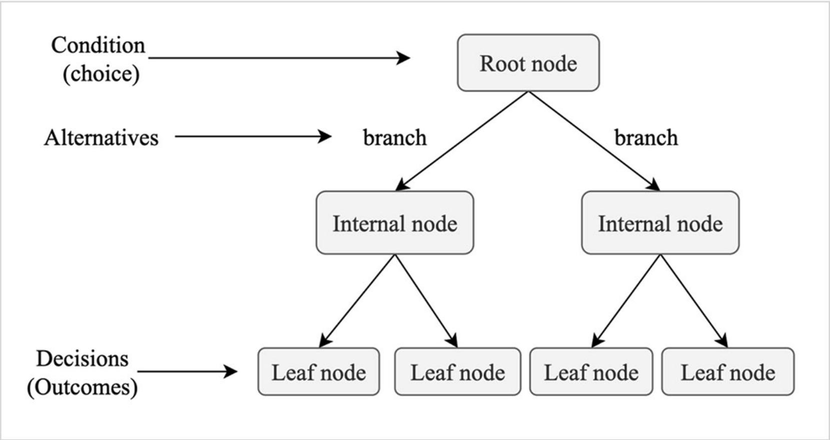

Fig. 1 Graphic structure of decision tree (DT)

architecture (Chaiken & Ledgerwood, 2012; Kaur et al., 2023). BP-NNs use interconnected neurons across multiple layers to process complex signals and adjust their internal parameters through backpropagation to minimize errors iteratively (Rumelhart et al., 1986). BP-NNs excel at processing complex signals, such as narrative texts and images, frequently encountered in crowdfunding (LeCun et al., 2015). We use both single-layer and two-layer BP-NN architectures in this study, where single-layer networks handle straightforward signals, and two-layer networks address more complex interactions. This architecture choice balances model complexity and generalization, mitigating overfitting risks, especially with limited data (Srivastava et al., 2014).

Our approach goes beyond traditional ML applications focused on prediction or variable extraction (Li et al., 2021; Ranta et al., 2023). By using DT and BP-NN models to emulate cognitive processing, we provide a nuanced understanding of investor behavior in high-noise environments, such as equity crowdfunding (Drover et al., 2018; Laureiro-Martínez & Brusoni, 2018). This method allows us to investigate how investors interpret different signal types under conditions of information asymmetry (Fig. 2).

4.3 Ensemble learning for enhanced prediction

Ensemble learning improves predictive accuracy by combining multiple base models (see Fig. 3). Techniques like random forest (RF) and gradient boosting (GB) enhance performance by aggregating predictions from diverse models (Breiman, 1996). RF generates independent base models through bootstrapped datasets, while GB builds models iteratively, focusing on rectifying errors from previous models. These methods serve as complements to DT and BP-NN, leveraging machine learning's computational power to enhance decision-making accuracy.

4.4 Prediction performance of ML models

We evaluate the performance of our ML models using four key metrics: accuracy, precision, recall, and $F1$ score, following established frameworks (Kaminski & Hopp, 2019). Accuracy measures the proportion of correct predictions. Precision indicates the proportion of true positives among predicted positives, while recall assesses the proportion of true positives among actual positives. The $F1$ score, a harmonic mean of precision and recall, provides a balanced metric to evaluate model performance, particularly when dealing with imbalanced datasets (Zhu et al., 2019). These metrics allow us to assess the effectiveness of DT and BP-NN models in predicting crowdfunding campaign success. In summary, our methodology employs DT and BP-NN models to simulate heuristic and systematic processing, respectively, and utilizes ensemble learning to enhance predictive performance. This approach provides a comprehensive framework for understanding how investors interpret and respond to diverse signals in the dynamic and high-noise environment of equity crowdfunding.

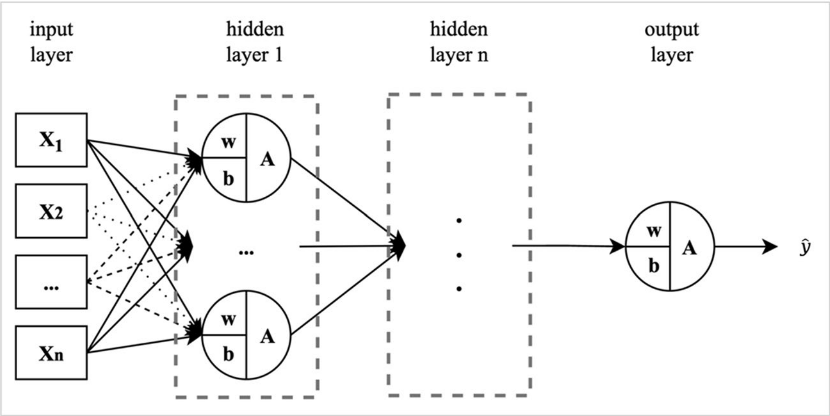

Fig. 2 A schematic of a backpropagation neural network. *Note*: full description of BP-NN model available in Rumelhart et al. (1986). w and b are the parameters to obtain the linear combination of input variables; then, the result will be transformed to the range between 0 and 1 through an activation function (A). $\hat{y}$ is the predicted value of the nth individual outcome

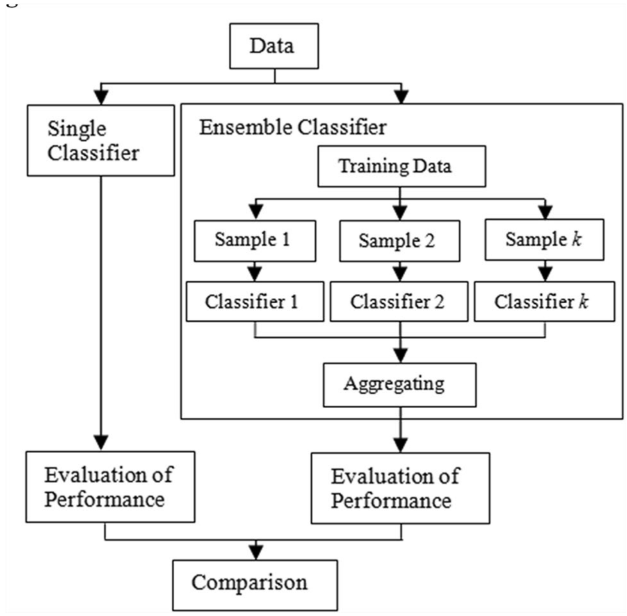

Fig. 3 Single vs. ensemble classifier. Note: full description of single vs. ensemble classifier available in Utami et al. (2014)

5 Data and variables

5.1 Data collection

We compile a unique dataset from the Dreammove platform, covering equity crowdfunding projects recorded between August 2014 and January 2019. The initial sample of 182 entries is systematically reduced to 144 through a rigorous data filtering process. Exclusions are made due to missing critical information: 23 entries lack actual investment amounts, 9 are missing team member photos, 4 do not include educational background information, and 2 lack project valuation data. This careful curation ensures the integrity and completeness of our final dataset, which forms a robust foundation for our analysis.

Dreammove is a key player in the Chinese equity crowdfunding market, making it a representative case for studying industry dynamics. Established in 2014 during a period of rapid market growth, Dreammove reflects the broader market's evolution, including the initial surge in platform numbers and the subsequent contraction driven by regulatory tightening and challenges such as poor project quality and immature investors. After manually reviewing the top 20 equity crowdfunding platforms by fundraising amount, we find that 14 official websites are now inactive, with Dreammove among the few that remain operational (see Appendix Table 13). This finding highlights Dreammove's resilience and adaptability in a volatile regulatory environment and underscores the scarcity of data sources in this sector.

The high rate of platform closures, reducing the number of active platforms from 186 in 2015 to just 23 by 2019, poses a significant challenge for expanding our sample size. With most equity crowdfunding platforms now inactive, it has become almost impossible to collect additional data to enlarge our sample. This constraint emphasizes the importance of Dreammove as one of the few remaining operational platforms that can provide valuable insights into the equity crowdfunding landscape in China.

Dreammove's project portfolio mirrors the industry distribution observed across 32 representative platforms, with a focus on Internet and IT services, as well as culture, sports, and leisure services, collectively representing about 50% of its projects. This alignment not only validates Dreammove's representativeness but also provides insights into the industries that dominate China's equity crowdfunding space (see Appendix Table 14). Located in Zhejiang, a major economic hub known for its vibrant private enterprise sector, Dreammove benefits from a region that actively fosters small and microsized business growth. Zhejiang's initiatives, such as the 'Three-Year Growth Plan for Small and Micro Enterprises,' highlight the province's commitment to private sector development. Dreammove's success in this environment underscores its role as a critical intermediary, connecting high-potential projects with investors, and exemplifying the broader trends in the Chinese equity crowdfunding market.

Dreammove's strong market performance further solidifies its significance. During its peak in 2015, it ranked 9th in market share among 113 equity crowdfunding platforms in China, commanding a notable 2.3% of the Internet private equity financing market. Its ability to navigate stringent regulatory measures and maintain a significant market presence underlines its robustness and adaptability, making it an ideal platform for this study.

5.2 Small sample size: challenges, justifications, and adaptive approaches

The dataset of 144 equity crowdfunding projects, while relatively small, is consistent with industry norms and existing research in the field. Studies like Eldridge et al. (2021) with 230 projects in the UK, Lukkarinen and Schwienbacher (2023) with 287 projects in Finland, and Prokop and Wang (2022) with 231 projects in Germany, as well as data from Beauhurst's annual reports, confirm that sample sizes in the low hundreds are typical for equity crowdfunding research. This consistency reinforces the representativeness of our dataset, despite its smaller size.

Nevertheless, small sample sizes present challenges, particularly in machine learning applications, where larger datasets are generally required for robustness and accuracy (LeCun et al., 2015). Smaller datasets increase the risk of overfitting, where models may perform well on training data but fail to generalize to new data (Srivastava et al., 2014). To address this limitation, we adopt several adaptive measures in the robustness check. First, we reduce feature dimensions, focusing on the most critical variables to minimize complexity and prevent overfitting. Additionally, we supplement our analysis with data from another platform to validate our findings, following best practices in research methodology. These strategies aim to enhance the reliability of our results, even within the constraints of a smaller dataset.

Our methodological adjustments and adherence to industry practices ensure that, despite the small sample size, our findings remain robust and meaningful. Future research can benefit from expanded data availability and improved model performance as the equity crowdfunding landscape continues to evolve.

5.3 Output variables

The primary outcome variable in our study is the success of equity crowdfunding campaigns, defined as a binary variable indicating whether a project reached its funding target within the campaign duration. This definition aligns with previous research (Calic & Mosakowski, 2016; Huang et al., 2021). We assign a code of 1 to projects that successfully secured funding equal to or exceeding their established funding goal, which applies to 92 out of the 144 projects in our dataset. This "all-or-nothing" model is standard in the crowdfunding industry, meaning investors are only obligated to contribute if the campaign meets its full funding target (Ralcheva & Roosenboom, 2019).

5.4 Input variables

Our study uses a comprehensive set of input variables, categorized into quantitative attributes, textual descriptions, and graphical images, to capture various aspects of the equity crowdfunding campaigns and the entrepreneurial teams behind them. Quantitative data, manually extracted from project statements, includes variables such as the project's target funding amount, starting bid, project valuation, equity allocation, share percentage held by the largest shareholder, CEO's shareholding status, quantitative data specific to the project, and the number of team members along with their educational background and industry experience that can be easily quantified without using sophisticated coding matrix (Mochkabadi & Volkmann, 2018).

To enhance our analysis, we create two novel variables using natural language processing (NLP) and image processing technology (IPT) available on the Baidu AI open platform (https://ai.baidu.com/), a widely used resource in academic research on Chinese business studies (Ruan et al., 2020; Yang et al., 2021). NLP helps extract implicit information contained within the textual descriptions of projects, while IPT facilitates the processing of information embedded in the photographic images of entrepreneurial teams.

Information extracted from the textual descriptions of projects is categorized into three distinct groups. Sentiment analysis (SA) focuses on detecting and analyzing the emotional tone conveyed within the text, including titles. Chinese segmentation (CS) dissects the text into its various linguistic components, such as verbs, nouns, and adjectives, while also considering the word count in the project title and introduction. Word vectorization (WV) assesses the grammatical quality of the texts and gauges their stylistic similarity to exemplary peer writings.

The information derived from the images of the management team is also organized into three distinct categories. Picture quality control (PQC) assesses the resolution of the photos. Face attribute analysis (FAA) employs the analysis of 150 facial attributes to calculate parameters such as beauty, age, and gender of team members. Emotion identification (EI) goes a step further to detect and capture the emotional expressions of individuals in the images. Additionally, we include the ratio of team members who have shared their photos in relation to the entire team. For detailed variable definitions and distributions, please refer to Table 1. All input variables are standardized in subsequent analyses to ensure consistency and enhance the predictive accuracy of our models.

6 Result discussion

6.1 Descriptive statistics

Table 2 provides an overview of our sample, presenting descriptive statistics and mean differences. The sample includes 144 equity crowdfunding projects, with a success rate of 64% (92 successful projects and 52 unsuccessful ones). We categorize the sample into two groups: successful vs. unsuccessful projects (columns (1) and (2)), and pre- vs. post-regulation projects (columns (4) and (5)). Columns (3) and (6) show mean differences between these groups, respectively.

In column (3), the most significant differences relate to quantitative information (panel A). Statistically significant mean differences are observed between successful and unsuccessful campaigns for variables such as *Sharetrans*, *Firstholder*, *CEO*, *Internet* and *Other*, implying that shareholding structure, CEO attributes and internet-compatibility are important to achieve crowdfund success. In panel B, statistically significant mean differences are reported for *Similar-class*, indicating that projects with higher industry-specific readability in their text descriptions are more likely to successfully secure fundings. In contrast, image information (panel C) shows less statistical significance.

However, in column (6), post-regulation changes reveal more pronounced differences in text and image information, suggesting a shift towards higher-quality qualitative data following the regulatory changes. Entrepreneurs are now likely to share more high-quality photos of their founding teams, characterized by better clarity, a more youthful and attractive appearance, and neutral emotional expressions. Text descriptions also show improved readability, a more neutral tone, and streamlined project introductions, with reduced length and less frequent use of verbs, nouns, and adjectives. These adjustments suggest an overall enhancement in qualitative information quality due to policy changes. Regarding quantitative data, it is notable that the largest shareholder has increased their ownership, while the number of founders with industry-specific experience has declined. It is important to emphasize that these observations are based on initial inter-group difference tests and do not account for the cognitive mechanisms of crowd investors.

Quantitative information remains the most reliable predictor of crowdfunding success; however, post-regulation improvements in text and image data indicate an increasing investor focus on qualitative aspects. This suggests while investors still prioritize quantitative data, there is a growing emphasis on the quality of qualitative information, reflecting a more holistic approach to evaluating campaigns. These changes underscore the evolving nature of investor decision-making, where both quantitative and qualitative signals are now being considered more critically in response to regulatory shifts and market maturation.

6.2 Prediction accuracy by signal types and combined portfolio

6.2.1 Quantitative vs. qualitative signals

Table 3 compares the prediction accuracy between quantitative and qualitative data. Regardless of the processing method employed, the results consistently show that models perform better with quantitative

Table 1 Variable definition and summary statistics

Variable name	Definition	Mean	SD	Min	Max
Panel A: quantitative					
Success	Whether the fundraising target achieved: 1 represents achieved	0.641	0.481	0	1
Target	Target amount of project (unit: 10,000 RMB)	122.5	115.0	2.400	1000
Startlimit	Minimum limitation of investment (unit: 10,000RMB)	0.755	0.870	0.070	10
Valuation	The logarithm of project valuation (unit: 10,000 RMB)	6.972	1.195	2.303	11.11
Sharetrans	Percentage of equity transferred (%)	12.78	11.20	0.010	57.14
Firstholder	share proportion of the largest shareholder	0.203	0.324	0	1
CEO	whether the CEO is the largest shareholder	0.387	0.489	0	1
Internet	Project type: based on internet	0.570	0.497	0	1
Technique	Project type: focus on technology	0.077	0.268	0	1
Market	Project type: expand market	0.930	0.257	0	1
Member	Number of team members	4.965	2.479	1	22
Entrepre	Average entrepreneurship experience	0.217	0.210	0	1
Industry	Average industry experience	0.783	0.313	0	1
Bachelor	Average education background of entrepreneurs: bachelor	0.432	0.343	0	1
Master	Average education background of entrepreneurs: master	0.131	0.211	0	1
Doctor	Average education background of entrepreneurs: doctor	0.022	0.076	0	0.429
MBA	Average education background of entrepreneurs: MBA	0.026	0.073	0	0.500
Abroad	Average education background of entrepreneurs: study abroad	0.050	0.109	0	0.500
Other	Number of numerals displaying in the project	0.944	1.184	0	7
Panel B: text					
Positive_t	Positive emotions in introduction calculated by sentiment analysis	0.887	0.317	0	1
Negative_t	Negative emotions in introduction calculated by sentiment analysis	0.070	0.257	0	1
Neutral_t	Neutral emotions in introduction calculated by sentiment analysis	0.042	0.202	0	1
Senti_title	Emotions implied in title: 2 (positive), 1 (neutral), 0 (negative)	1.599	0.547	0	2
Length_title	Number of words used in title	9.634	5.944	2	25
Length_intro	Number of words used in introduction	63.86	24.14	9	100
Verb	Number of verbs used in introduction	9.028	4.466	0	24
Noum	Number of nouns used in introduction	11.30	4.874	0	26
Adjv	Number of adjectives used in introduction	2.282	2.142	0	11
Readability	Grammatically well-written score of introductions based on DNN model	6.364	0.678	4.345	8.227
Similar_fund	Similarity to the best peer writing with the highest fundraising percentage	0.264	0.118	0	0.546
Similar_class	Similarity to the best peer writing in the same business scope	0.238	0.206	0	1
Similar_all	Similarity to the best peer writing with the most fundraising	0.265	0.150	0	1
Panel C: image					
Clarity	Average resolution score of pictures	0.100	0.195	0	0.987
Youth	Average age of entrepreneurs among pictures	21.13	11.13	0	39.67
Beauty	Average beauty score of entrepreneurs among pictures	38.71	21.11	0	77.42
Male	Male percentage of entrepreneurs among pictures	0.462	0.439	0	1
Happy	Average score of happy sentiment among pictures	0.282	0.366	0	1
Fear	Average score of fear sentiment among pictures	0.002	0.028	0	0.333
Neutral_i	Average score of neutral sentiment among pictures	0.162	0.283	0	1
NumPho	Average degree of photo disclosure within a project	0.761	0.378	0	1

Table 2 Descriptive statistics and mean difference

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Fail ($N=52$)	Success ($N=92$)	Mean Diff	Before regulation ($N=95$)	After regulation ($N=49$)	Mean Diff
Panel A: quantitative						
Target	−0.046	0.026	−0.071	0.083	−0.162	0.245
Startlimit	−0.108	0.061	−0.170	0.011	−0.022	0.034
Valuation	0.165	−0.093	0.258	0.024	−0.047	0.071
Sharetrans	−0.216	0.122	−0.337*	0.049	−0.095	0.144
Firstholder	−0.262	0.148	−0.411**	−0.272	0.528	−0.800***
CEO	0.250	0.457	−0.207**	0.337	0.469	−0.133
Internet	0.462	0.641	−0.180**	0.547	0.633	−0.085
Technique	0.115	0.054	0.061	0.084	0.061	0.023
Market	0.904	0.946	−0.042	0.926	0.939	−0.012
Member	0.096	−0.054	0.150	0.070	−0.136	0.206
Entrepre	0.175	−0.099	0.275	0.004	−0.007	0.011
Industry	0.123	−0.069	0.192	0.193	−0.375	0.568***
Bachelor	−0.109	0.062	−0.171	−0.097	0.188	−0.284
Master	0.113	−0.064	0.177	0.034	−0.067	0.101
Doctor	−0.060	0.034	−0.094	0.080	−0.155	0.235
MBA	−0.170	0.096	−0.266	0.070	−0.135	0.205
Abroad	−0.114	0.064	−0.179	0.072	−0.140	0.213
Other	−0.186	0.105	− −0.290*	0.050	−0.097	0.147
Panel B: text						
Positive_t	0.923	0.870	0.054	0.926	0.816	0.110**
Negative_t	0.038	0.087	−0.048	0.053	0.102	−0.049
Neutral_t	0.038	0.043	−0.005	0.021	0.082	−0.061*
Senti_title	0.091	−0.052	0.143	0.070	−0.135	0.205
Length_title	0.129	−0.073	0.202	0.030	−0.059	0.089
Length_intro	−0.084	0.047	−0.131	0.186	−0.361	0.548***
Verb	−0.124	0.070	−0.195	0.106	−0.205	0.311*
Noum	−0.037	0.021	−0.058	0.121	−0.234	0.354**
Adjv	0.089	−0.050	0.139	0.125	−0.242	0.367**
Readability	0.137	−0.077	0.214	−0.126	0.243	−0.369**
Similar_fund	−0.039	0.022	−0.062	−0.079	0.153	−0.232
Similar_class	−0.241	0.136	−0.377**	0.065	−0.125	0.190
Similar_all	0.012	−0.007	0.019	−0.105	−0.043	0.066
Panel C: image						
Clarity	0.142	−0.080	0.223	−0.105	0.203	−0.308*
Youth	0.176	−0.099	0.275	−0.152	0.295	−0.448**
Beauty	−0.005	0.003	−0.008	−0.204	0.395	−0.599***
Male	−0.050	0.0280	−0.079	−0.033	0.063	−0.096
Happy	−0.011	0.006	−0.018	0.085	−0.165	0.250
Fear	−0.083	0.047	−0.130	0.043	−0.083	0.126
Neutral_i	0.082	−0.047	0.129	−0.267	0.517	−0.784***
NumPho	0.085	−0.048	0.133	−0.209	0.405	−0.613***

Note: *Significant at 10%

**Significant at 5%

***Significance at 1%

data. Decision tree (DT) models, which simulate heuristic, rule-based decision-making, achieve the highest accuracy of 0.66 with quantitative data, but their accuracy drops to 0.48 when applied to qualitative text. Similarly, backpropagation neural network (BP-NN) models, which involve more complex cognitive processing, also show higher accuracy with quantitative data (0.59 for both BP-NN (1) and BP-NN (2)), while their performance with qualitative data decreases to 0.48.

These findings offer compelling evidence in support of H1a, decisively showing that quantitative data holds greater predictive power than qualitative data in determining crowdfunding outcomes. Across all models and processing methods, the results consistently demonstrate that investors prioritize quantitative metrics, such as shareholding structure and CEO attributes, over qualitative aspects like text descriptions and images. This clear preference highlights the critical role of concrete, measurable information in driving investment decisions, reinforcing the notion that investors rely more heavily on quantitative signals when evaluating potential success.

However, many studies on machine learning models for outcome prediction, when comparing different data types or models, tend to focus solely on evaluating accuracy metrics, while bypassing formal statistical tests (Athey & Imbens, 2019; Gu et al., 2020; Mullainathan & Spiess, 2017), potentially leading to less reliable findings. To ensure robust comparisons, we conduct formal tests to rigorously assess performance differences between quantitative and qualitative signals. We run models for both data types multiple times, each using a different random seed to control key processes such as data splitting, shuffling, and weight initialization, which introduces natural variability into the model (Pineau et al., 2021). This approach captures a wider range of outcomes and prevents reliance on any single configuration. Specifically, we randomly selected 10 seeds based on the model in Table 3. From these runs, we generated 10 sets of performance metrics—accuracy, precision, recall, and $F1$ score—for both quantitative and qualitative signals. To assess the statistical significance of the performance differences, we conducted t-tests on the means of these metrics. Table 4 presents the

Table 3 Classification results by quantitative and qualitative signals

Classifier/data		Acc	Prec	Rec	$F1$
Panel A: signal type—quantitative					
Base learner	DT	0.66	0.78	0.66	0.54
	BP-NN (1)	0.59	0.57	0.59	0.58
	BP-NN (2)	0.59	0.57	0.59	0.58
Ensemble learning	RF-DT	0.64	0.62	0.64	0.62
	GB-DT	0.61	0.59	0.61	0.59
	GB-BP (1)	0.59	0.57	0.59	0.58
	GB-BP (2)	0.59	0.57	0.59	0.58
Panel B: signal type—qualitative					
Base learner	DT	0.48	0.43	0.48	0.45
	BP-NN (1)	0.48	0.45	0.48	0.46
	BP-NN (2)	0.48	0.45	0.48	0.46
Ensemble learning	RF-DT	0.55	0.50	0.55	0.52
	GB-DT	0.45	0.42	0.45	0.43
	GB-BP (1)	0.55	0.50	0.55	0.52
	GB-BP (2)	0.55	0.50	0.55	0.52

Note: DT decision tree, *BP-NN* backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting. DT models represent rule-of-thumb processing, which requires minimal cognitive effort based on existing knowledge, while BP-NN models represent reflective learning, which demands greater cognitive effort. Specifically, BP-NN (1) simulates linear relationships using a single-layer neural network, while BP-NN (2) addresses more complex non-linear relationships with a two-layer network. Ensemble methods applied to both DT and BP-NN models *t*-test results (panel C) alongside the average performance metrics for each group—quantitative in panel A and qualitative in panel B. The results indicate that the mean performance metrics for quantitative signals (panel A) significantly outperform those for qualitative signals (panel B). Furthermore, panel C reveals that the performance differences between quantitative and qualitative signals are statistically significant.

6.2.2 *Prediction accuracy across individual and combined signal types*

Table 5 provides a detailed comparison of prediction accuracy for segmental (individual signal types) and holistic (combined signal bundle) analysis. Panel A reports the accuracy of models based solely on quantitative data, reflecting similar results to those

Table 4 Formal comparison tests between quantitative and qualitative signals

Classifier/data		Acc	Prec	Rec	F1
Panel A: signal type—quantitative					
Base learner	DT	0.624	0.552	0.624	0.548
	BP-NN (1)	0.619	0.606	0.619	0.608
	BP-NN (2)	0.607	0.599	0.607	0.601
Ensemble learning	RF-DT	0.642	0.630	0.642	0.626
	GB-DT	0.650	0.638	0.650	0.637
	GB-BP (1)	0.613	0.591	0.613	0.599
	GB-BP (2)	0.618	0.605	0.618	0.609
Panel B: signal type—qualitative					
Base learner	DT	0.567	0.492	0.567	0.520
	BP-NN (1)	0.573	0.559	0.573	0.560
	BP-NN (2)	0.548	0.540	0.548	0.543
Ensemble learning	RF-DT	0.583	0.540	0.583	0.543
	GB-DT	0.551	0.552	0.551	0.540
	GB-BP (1)	0.568	0.543	0.568	0.552
	GB-BP (2)	0.565	0.544	0.565	0.550
Panel C: difference (quantitative – qualitative)					
Base learner	DT	0.057**	0.060	0.057**	0.028
	BP-NN (1)	0.046**	0.047**	0.046**	0.048**
	BP-NN (2)	0.059***	0.059***	0.059***	0.058***
Ensemble learning	RF-DT	0.059**	0.090**	0.059**	0.083***
	GB-DT	0.099**	0.086***	0.099**	0.097**
	GB-BP (1)	0.045***	0.048**	0.045***	0.047**
	GB-BP (2)	0.053**	0.061**	0.053**	0.059**

Note: *DT* decision tree, *BP-NN* backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting. DT models represent rule-of-thumb processing, which requires minimal cognitive effort based on existing knowledge, while BP-NN models represent reflective learning, which demands greater cognitive effort. Specifically, BP-NN (1) simulates linear relationships using a single-layer neural network, while BP-NN (2) addresses more complex non-linear relationships with a two-layer network. Ensemble methods applied to both DT and BP-NN models. Panel A and Panel B report the means of the four prediction metrics for quantitative information and qualitative information, respectively, across the 10 random samples. Panel C reports the differences in the means of the four prediction metrics between quantitative information and qualitative information across the 10 random samples, along with their significance

***Statistical significance at the 1% level

**Statistical significance at the 5% level

*Statistical significance at the 10% level in Table 3. Panels B and C split qualitative information into text and image data, assessing their predictive accuracy using various machine learning (ML) models. Panel D presents the accuracy of models that integrate all three data types into a combined signal portfolio.

The table reveals that models using individual data types consistently outperform the combined signal portfolio. Quantitative data continues to deliver the highest accuracy, with the decision tree (DT) model achieving a prediction accuracy of 0.66. Text-based models also perform well, with the gradient boosting BP-NN (1) achieving the highest accuracy at 0.68. Image data, however, produces more moderate results, with BP-NN (2) achieving a maximum accuracy of 0.59. These findings further support H1a, confirming that quantitative signals remain the most reliable predictor of campaign success.

However, when signals are combined into a bundle, prediction accuracy does not improve significantly. In fact, the combined models underperform compared to standalone quantitative models. Even with the highest accuracy in the combined signal portfolio produced by the RF-DT model, the expected marginal improvement over standalone quantitative data is not observed. This suggests that integrating multiple signals introduces complexities and conflicts, diminishing rather than enhancing predictive performance.

The results strongly support H1b, indicating that the predictive accuracy of a signal portfolio depends on the interplay between quantitative and qualitative signals. Synergy between signals enhances campaign success, while conflicting signals reduce effectiveness. These findings highlight the challenges investors face when synthesizing multiple signal types, particularly in environments prone to information overload. The results underscore the need for clearer, more cohesive signal presentations to improve investor decision-making.

6.2.3 *Signal processing and investor decision-making*

Although entrepreneurs present bundled signals simultaneously, crowd investors may not process them all at once. Due to limited attention and cognitive constraints, signals are often assessed separately, requiring investors to reconfigure their mental focus when switching between tasks (Thiele et al., 2022). This section explores how different signal types are processed within the dual-system model.

Quantitative information is critical for investors in distinguishing high-quality projects, as it can be evaluated using common sense and domain knowledge accumulated through experience. This pre-existing knowledge simplifies the processing, making quantitative analysis relatively straightforward. As a result, quantitative data remains the most valuable signal type, with decision tree (DT) models consistently outperforming BP-NN models in terms of accuracy, precision, and recall (Table 5, panel A). These results support H2a, suggesting that investors rely on quick, rule-based processing when evaluating quantitative information, which explains the effectiveness of DT models. Interestingly, ensemble classifiers such as random forest (RF) and gradient boosting (GB) do not consistently improve the performance of standalone DT or BP-NN models. This suggests that ensemble methods may not be as effective in reducing variance associated with quantitative data as expected.

Textual information introduces more complexity compared to quantitative data. Entrepreneurs can embellish descriptions, creating ambiguity and requiring more effortful, systematic processing. Table 5 panel B shows that BP-NN models consistently outperform DT models across all metrics, underscoring the need for deeper cognitive effort to analyze text. Notably, the single-layer BP-NN model slightly outperforms the two-layer BP-NN, indicating that a simpler, linear approach is more effective than a more complex, non-linear one when interpreting narrative texts. These findings support H2b, demonstrating that narrative texts, with their inherent ambiguity and complexity, demand analytical processing for accurate predictions of campaign success. The superior performance of BP-NN models highlights the importance of systematic processing, while the edge of the simpler single-layer BP-NN model suggests that while analytical effort is crucial, a balanced and efficient approach is optimal. Ensemble classifiers—RF-DT, GB-BP-NN (1), and GB-BP-NN (2)—generally improve the performance of both DT and BP-NN models. However, GB-DT fails to enhance the DT model, suggesting that ensemble methods, while designed to reduce bias and variance, do not always succeed due to variations in dataset characteristics.

Table 5 Classification results by signal types and combined portfolio

Classifier/data		Acc	Prec	Rec	F1
Panel A: signal type—numerical					
Base learner	DT	0.66	0.78	0.66	0.54
	BP-NN (1)	0.59	0.57	0.59	0.58
	BP-NN (2)	0.59	0.57	0.59	0.58
Ensemble learning	RF-DT	0.64	0.62	0.64	0.62
	GB-DT	0.61	0.59	0.61	0.59
	GB-BP (1)	0.59	0.57	0.59	0.58
	GB-BP (2)	0.59	0.57	0.59	0.58
Panel B: signal type—text					
Base learner	DT	0.64	0.40	0.64	0.49
	BP-NN (1)	0.66	0.64	0.66	0.62
	BP-NN (2)	0.64	0.61	0.64	0.60
Ensemble learning	RF-DT	0.66	0.64	0.66	0.62
	GB-DT	0.59	0.61	0.59	0.60
	GB-BP (1)	0.68	0.68	0.68	0.63
	GB-BP (2)	0.66	0.64	0.66	0.62
Panel C: signal type—image					
Base learner	DT	0.55	0.50	0.55	0.52
	BP-NN (1)	0.57	0.46	0.57	0.49
	BP-NN (2)	0.59	0.52	0.59	0.53
Ensemble learning	RF-DT	0.50	0.47	0.50	0.48
	GB-DT	0.52	0.52	0.52	0.52
	GB-BP (1)	0.59	0.52	0.59	0.53
	GB-BP (2)	0.59	0.52	0.59	0.53
Panel D: signal portfolio					
Base learner	DT	0.55	0.53	0.55	0.54
	BP-NN (1)	0.57	0.55	0.57	0.56
	BP-NN (2)	0.55	0.53	0.55	0.54
Ensemble learning	RF-DT	0.64	0.61	0.64	0.60
	GB-DT	0.59	0.57	0.59	0.58
	GB-BP (1)	0.57	0.56	0.57	0.57
	GB-BP (2)	0.59	0.58	0.59	0.58

Note: *DT* decision tree, *BP-NN* backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting. DT models represent rule-of-thumb processing, which requires minimal cognitive effort based on existing knowledge, while BP-NN models represent reflective learning, which demands greater cognitive effort. Specifically, BP-NN (1) simulates linear relationships using a single-layer neural network, while BP-NN (2) addresses more complex non-linear relationships with a two-layer network. Ensemble methods applied to both DT and BP-NN models

Table 5 panel C reveals that BP-NN models outperform DT models, indicating that image analysis requires more reflective cognitive processing. BP-NN models are better equipped to capture complex, non-linear patterns, whereas DT relies on simpler, rule-based processing. The superior performance of BP-NN (2) over BP-NN (1) suggests that interpreting visual cues, such as facial expressions, demands higher-order cognitive effort, further supporting H2b. Inconsistent results from ensemble classifiers indicate that these techniques do not always enhance performance. RF-DT, GB-DT, and GB-BP-NN (2) fail to improve their base models, while GB-BP-NN (1) enhances BP-NN (1)'s performance. This suggests that the effectiveness of ensemble methods depends on the data and techniques used. Each signal type requires different cognitive processing approaches, with quantitative data benefiting from heuristic processing and qualitative data requiring systematic analysis.

Despite the conflicting nature of signals within the combined portfolio, Table 5 panel D reveals subtle insights. While both DT and BP-NN models generally perform worse in predicting success compared to individual signals, some differences emerge. Notably, the single-layer BP-NN (1) model achieves slightly higher prediction accuracy (0.57) compared to the DT model (0.55) and the more complex BP-NN (2) model (0.55). This pattern is consistent across other metrics, highlighting nuanced differences between the models. This suggests that a simpler, linear processing approach, as represented by BP-NN (1), may be more effective for integrating diverse signals than the more complex BP-NN (2). These findings support H2c, which posits that holistic processing in a high-noise crowdfunding environment is likely to trigger more intuitive, less effortful processing, leading to more efficient predictions of campaign success. Conversely, while the DT model excels at making quick, rule-based decisions with minimal cognitive effort, it struggles with holistic processing, which requires synthesizing diverse and complex information—a task demanding higher-order cognitive strategies. This further supports H2c, showing that while both the DT model and BP-NN (1) rely on heuristic and intuitive processing, BP-NN (1)'s more reflective approach aligns better with the demands of holistic processing.

In summary, while simpler holistic processing shows some effectiveness, neither increased complexity (as seen in BP-NN (2)) nor rule-based approaches (as seen in DT) consistently improve outcomes in this context.

6.3 Cross-validation tests

Cross-validation enhances model evaluation by using multiple data splits, which reduces overfitting, optimizes hyperparameters, and improves generalizability (Arlot & Celisse, 2010; Athey & Imbens, 2019; Stone, 1974). It also ensures more reliable comparisons by using consistent datasets across runs, increasing confidence in the results (Ranta et al., 2023).

In our study, we align with these best practices by implementing cross-validation to ensure accuracy and reliability in our evaluations. We further refine our methodology with two key parameter adjustments to enhance sample splitting. First, we change the random seed from 0 to 1 during sample splitting. The random seed determines the algorithm's sampling rules and, while its value can be arbitrary, predefining it ensures repeatability in our results (Pineau et al., 2021). This adjustment allows for consistent data partitioning across multiple runs, crucial for validating our findings. Panel A in Table 6 shows the results after this adjustment. Second, we modify the training-test split ratio. Our primary analysis uses a 70%/30% split, balancing a sufficiently large training sample with a representative test sample (Hastie et al., 2009). This ratio is widely used in practice and provides a solid foundation for model evaluation. We also test alternative splits, including 80%/20% (panel B) and 75%/25% (panel C). Our results show that these adjustments maintain the robustness of our findings, further reinforcing the validity of our conclusions.

6.4 Interpreting machine learning models: marginal impact analysis using SHAP

Machine learning models are often seen as "black boxes" due to the difficulty in interpreting their predictions and estimating marginal impacts, a challenge not present in traditional regression models (Hassija et al., 2024; Rai, 2020). This lack of transparency raises concerns about the reliability and trustworthiness of the model's predictions, as users may struggle to understand how specific input features influence outcomes.

To address these challenges, we apply the SHAP (SHapley Additive exPlanations) method, a powerful tool for interpreting machine learning models and clarifying the contributions of different data types within our framework (Erel et al., 2021). SHAP assigns a value to each feature based on its contribution to the model's output, offering a clear view of how various data types interact and influence the decision-making process. By leveraging SHAP, we bridge the gap between the inherent complexity of machine learning models and the need for interpretability.

Our approach involves two steps: step one—dimensionality reduction. We first apply principal component analysis (PCA) to each signal type: quantitative, narrative, and visual. This is especially relevant in the crowdfunding context, where investors face information overload and conflicting signals (Block et al., 2018). PCA helps reduce noise and enhances the signal-to-noise ratio by distilling each signal type into its most significant components (Jolliffe & Cadima, 2016; Plummer et al., 2016; Vismara, 2018).

Table 6 Cross-validation test: classification results of different parameters

Classifier/data		Acc	Prec	Rec	F1
Panel A: change random seed from 0 to 1					
Numerical	DT	0.79	0.81	0.79	0.77
	BP-NN (1)	0.67	0.70	0.67	0.68
	BP-NN (2)	0.72	0.74	0.72	0.73
Text	DT	0.63	0.62	0.63	0.62
	BP-NN (1)	0.72	0.72	0.72	0.69
	BP-NN (2)	0.72	0.72	0.72	0.69
Image	DT	0.67	0.65	0.67	0.65
	BP-NN (1)	0.72	0.71	0.72	0.70
	BP-NN (2)	0.58	0.57	0.58	0.57
Bundled	DT	0.49	0.66	0.49	0.46
	BP-NN (1)	0.63	0.65	0.63	0.64
	BP-NN (2)	0.58	0.62	0.58	0.59
Panel B: training/test set (0.8/0.2)					
Numerical	DT	0.66	0.43	0.66	0.52
	BP-NN (1)	0.62	0.63	0.62	0.62
	BP-NN (2)	0.62	0.63	0.62	0.62
Text	DT	0.52	0.54	0.52	0.53
	BP-NN (1)	0.66	0.62	0.66	0.60
	BP-NN (2)	0.66	0.62	0.66	0.60
Image	DT	0.55	0.40	0.55	0.47
	BP-NN (1)	0.62	0.54	0.62	0.55
	BP-NN (2)	0.66	0.62	0.66	0.60
Bundled	DT	0.52	0.39	0.52	0.45
	BP-NN (1)	0.66	0.66	0.66	0.66
	BP-NN (2)	0.66	0.66	0.66	0.66
Panel C: training/test set (0.75/0.25)					
Quantitative	DT	0.64	0.41	0.64	0.50
	BP-NN (1)	0.61	0.60	0.61	0.60
	BP-NN (2)	0.58	0.58	0.58	0.58
Text	DT	0.64	0.63	0.64	0.64
	BP-NN (1)	0.67	0.65	0.67	0.62
	BP-NN (2)	0.67	0.65	0.67	0.62
Image	DT	0.50	0.46	0.50	0.47
	BP-NN (1)	0.61	0.56	0.61	0.55
	BP-NN (2)	0.64	0.61	0.64	0.60
Bundled	DT	0.50	0.43	0.50	0.45
	BP-NN (1)	0.58	0.59	0.58	0.59
	BP-NN (2)	0.61	0.62	0.61	0.62

Note: *DT* decision tree, *BP-NN* Backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting

For quantitative data (panel A, Table 7), six principal components are identified: Fn1 highlights educational background, Fn2 covers project characteristics, Fn3 addresses financial aspects, Fn4 focuses on ownership structure, Fn5 merges equity dynamics with educational credentials, and Fn6 emphasizes entrepreneurs' experience. These components reveal how project and team attributes impact equity crowdfunding success.[3] For narrative text data (panel B, Table 7), five principal components emerged: Ft1 captures linguistic complexity, Ft2 emphasizes emotional tone, Ft3 reflects textual similarity to successful peers, Ft4 focuses on neutrality and project alignment, and Ft5 highlights emotional tone in titles and descriptions. These components illustrate how linguistic features and emotional tone affect crowdfunding presentation and credibility.[4] For image data (panel C, Table 7), three principal components are identified: Fi1 examines visual and demographic appeal, Fi2 focuses on visual emotional tone, and Fi3 addresses image quality and negative cues. These components highlight the importance of visual and emotional

[3] Fn1: variables like "Bachelor," "Master," "Doctor," "Abroad," and "MBA" have the highest positive loadings, suggesting that Fn1 is heavily influenced by educational and international experience factors. Fn2: "Internet" and "Member" have strong positive loadings, while "Technique" has a strong negative loading, indicating that this component might reflect a contrast between digital presence and technical aspects. Fn3: "Valuation" and "Target" show high positive loadings, suggesting that this component may be driven by company valuation metrics. Fn4: "CEO" has a strong positive loading, indicating this component is significantly influenced by executive leadership. Fn5: "Sharetrans" has a high negative loading, suggesting that this component might be associated with share transaction dynamics. Fn6: "Entrepre" and "Industry" have strong positive loadings, hinting that this component is likely linked to entrepreneurial and industry-related factors.

[4] Ft1: high positive loadings for Verb, Noun, and Length_intro suggest this component is driven by syntactic structure and introductory content. Ft2: strong positive loading for Negative_t and negative loading for Positive_t indicate this component captures sentiment polarity, focusing on the distinction between positive and negative tones. Ft3: positive loadings for Similar_c ~ s and Similar_all, and a negative loading for DNN, imply this component reflects content similarity and contrasts with deep neural network features. Ft4: negative loadings for Neutral_t, Adjective, and Number suggest this component is influenced by neutrality and descriptive elements. Ft5: high positive loading for Senti_title, along with positive contributions from Number and Length_intro, indicate this component reflects the influence of title sentiment and content length.

Table 7 PCA results for quantitative, text, and image variables

Panel A: numerical	fn1	fn2	fn3	fn4	fn5	fn6
Target	−0.0529	0.1092	0.612	0.0502	−0.0574	−0.0466
Startlimit	−0.2207	0.1791	0.0231	0.0771	0.1208	0.1899
Valuation	0.0178	−0.0772	0.6301	−0.0569	0.0546	0.0118
Sharetrans	0.0135	0.1826	−0.0204	0.0284	−0.6407	−0.0548
Member	0.0828	0.0715	−0.2285	−0.0572	0.2848	−0.0133
Firstholder	−0.0397	0.0624	0.0157	0.686	−0.0748	−0.0363
CEO	0.0367	−0.0579	−0.0144	0.6963	0.0689	0.0335
Entrepre	0.0471	−0.0419	−0.0811	0.0322	−0.1395	0.6449
Industry	−0.0513	0.0391	0.0713	−0.033	0.1336	0.6601
Bachelor	0.4326	0.1068	0.0897	0.0874	−0.2328	−0.1375
Master	0.5355	0.0042	0.0564	−0.0069	−0.0687	0.177
Doctor	0.4692	−0.0356	−0.1095	−0.0816	0.2762	−0.1345
MBA	0.1858	−0.1027	0.3647	0.0487	0.0498	0.1706
Abroad	0.4523	−0.0351	−0.0734	0.031	0.0205	−0.0244
Technique	−0.0064	−0.6317	0.0276	0.0495	−0.0059	0.0059
Market	−0.003	0.625	0.0262	0.0423	−0.0431	0.0128
Internet	−0.0144	0.295	0.0128	0.0689	0.5476	−0.0787
Panel B: text	ft1	ft2	ft3	ft4	ft5	
Positive_t	0.0058	−0.5733	0.0246	−0.3228	0.0283	
Negative_t	0.0035	0.7426	0.0198	−0.1507	0.0114	
Neutral_t	−0.0136	−0.0433	−0.0638	0.6991	−0.0589	
Noum	0.4796	−0.045	−0.0398	0.0004	0.0447	
Verb	0.5135	−0.0035	−0.0013	−0.0041	−0.0918	
Adjv	0.2187	0.1607	−0.233	−0.3514	−0.1051	
Number	0.1763	0.0417	0.2427	0.4413	0.2654	
Length_intro	0.5487	0.0139	0.0067	−0.0027	0.0693	
DNN	0.1314	−0.1399	−0.32	0.1052	−0.2774	
Similiar_class	−0.1077	0.1474	0.5878	−0.124	0.0773	
Similiar_fund	0.2867	0.0803	0.2807	0.1214	−0.1755	
Similiar_all	0.1099	−0.188	0.5717	0.1069	−0.1221	
Length_title	−0.0239	−0.0427	−0.094	−0.0679	0.6659	
Senti_title	0.0489	0.0573	0.1335	0.0947	0.5717	
Panel C: image	fi1	fi2	fi3			
Photo	0.512	−0.0235	0.0577			

Table 7 (continued)

Clarity	0.0787	−0.0746	0.6474
Beauty	0.4846	0.0906	0.0062
Youth	0.496	−0.1061	0.0083
Male	0.387	0.1345	−0.1523
Fear	−0.0503	0.0506	0.7407
Happy	0.2272	−0.6819	−0.0547
Neutral_i	0.2167	0.6991	−0.0514

Note: using eigenvalues greater than 1 as the selection criterion, 6 principal components were extracted for quantitative information (panel A), 5 for text information (panel B), and 3 for image information (panel C). Table 15 presents the scoring coefficients for each principal component across these three information categories. The sum of squares (column-loading) equals 1

elements in shaping investor perception of crowdfunding projects.[5]

Step two—SHAP Analysis: we then employ SHAP to distinguish the contributions of each PCA component from different signal types. This technique quantifies the impact of each component on the model's predictions, providing a clearer and more integrated understanding of how these signals collectively influence outcomes, as recommended by Erel et al. (2021). SHAP assigns an importance value (SHAP value) to each feature based on its contribution to the model's predictions, making it particularly effective for interpreting complex models like neural networks (Lundberg & Lee, 2017).

The SHAP plot (Fig. 4), displayed as a bar chart, highlights the mean absolute SHAP values to emphasize the importance of each component. The analysis reveals that Fn4 (founder ownership) and Ft1 (linguistic complexity) are the most critical features, with Fn4 showing a significant negative impact and Ft1 demonstrating a strong positive effect. Fn4, representing founder ownership, negatively impacts predictions, possibly signaling rigidity or a reluctance to accept external input, which could detract from perceived growth potential. Additionally, Fn5's negative association with share transaction dynamics suggests that reduced equity transfers to crowd investors are perceived positively, signaling stability and long-term commitment from founders. This preference for steady ownership over frequent changes likely reflects investor concerns about instability.

Overall, the Fn components exhibit strong positive contributions (+ 0.098), offset by notable negative effects (− 0.069). The Ft components show the lowest positive impact (+ 0.078) but the highest negative influence (− 0.104), indicating that textual information exerts a more volatile influence. In contrast, the

[5] Fi1: high positive loadings for Photo, Beauty, and Youth suggest this component is driven by visual appeal and youthfulness. Male, Happy, and Neutral_i also contribute positively, while Fear has a slight negative influence. Fi2: Neutral_i strongly influences this component, with Happy showing a strong negative loading, indicating a contrast between neutrality and positive emotions. Clarity also negatively contributes. Fi3: Fear and Clarity have strong positive loadings, making them key drivers of this component, while Male shows a negative influence.

Fi components provide consistent positive contributions (+0.081), reflecting the importance of visual signals in investor decision-making.

In summary, the analysis shows that quantitative signals (Fn components) are the most influential in predicting crowdfunding success. Investors prioritize structured quantitative data, such as equity dynamics and project attributes, which align with their preference for clarity and stability. Narrative texts (Ft components) have a more volatile impact. While linguistic complexity (Ft1) positively influences outcomes, unoriginal narratives (Ft3) and overly emotional tones in titles and descriptions (Ft5) and reduce credibility. This suggests that while narratives can enhance a campaign, they must be crafted carefully to avoid undermining investor confidence. Visual signals (Fi components) also play a significant role, particularly Fi1 (visual and demographic appeal). Fi1 has high positive loadings for beauty and youth suggest visual appeal and youthfulness can positively impress investors and lead to higher funding success, suggesting that visually appealing images are processed more efficiently through intuitive heuristics, leading to more accurate predictions of campaign success.

For entrepreneurs, these findings indicate that clear, stable quantitative information, coupled with a balanced narrative and authentic visuals, is most effective in signaling quality to investors. Traditional indicators like ownership structure or imitating successful peers may have unintended negative effects, this highlights the complex ways in which different data components influence model predictions, reflecting deeper investor concerns about authenticity, innovation, and long-term stability.

6.5 Policy effects on distinctive signal types and combined portfolio

Equity crowdfunding in China faces challenges like poor project quality, unclear platform roles, and inadequate supervision, echoing issues from the P2P lending crisis. To address these, the Chinese government introduced the 2016 *Interim Measures*, aimed at enhancing transparency and investor protection. Although originally designed for P2P lending, these regulations have significantly influenced equity crowdfunding by mandating comprehensive disclosures and robust risk management practices, thus improving signal credibility and reducing information asymmetry (Ahlers et al., 2015; Vismara, 2016). To assess the policy impact, we divide our sample into pre- and post-regulation periods to evaluate whether stricter disclosure requirements enhanced the predictive accuracy of different types of information using machine learning models.

6.5.1 Policy impact on signal quality

The 2016 policy requires founders to disclose detailed project information, improving signal quality. Table 8 shows that post-regulation, quantitative signals such as Firstholder, CEO attributes, and Internet presence became stronger predictors of success. This suggests that these indicators became clearer and more reliable, supporting the notion that investors value structured quantitative data. Text-based signals also see improvements, with enhanced readability and content alignment post-regulation. Variables like title length and verb usage become more predictive, making textual content a more effective tool for signaling quality. Similarly, visual signals show improvements in clarity and portrayal of team members, although male representation in images become negatively associated, indicating changing perceptions among investors. Overall, these regulatory measures have significantly enhanced signal quality, making them more reliable for predicting campaign success and helping investors make more informed decisions.

6.5.2 Impact on prediction accuracy

Table 9 compares prediction results before and after regulatory changes for different signal types and combined signal bundles. For quantitative information (panel A), post-regulation, all models show a 14% average increase in accuracy. It is important to note that the quantitative signals remain largely similar after the regulatory intervention, as indicated in Table 3. Drawing on the theory of investor attention, we suggest that effective noise control in crowdfunding post-regulation enables investors to allocate more focus to quantitative information. This shift in attention contributes to the observed enhancement in accuracy. GB-DT consistently outperforms other models, highlighting its robustness in handling structured data.

For textual information (panel B), the regulatory impact on textual data is mixed. While DT, BP-NN (2), and RF-DT models show improved accuracy, BP-NN (1) and GB-BP-NN models perform worse.

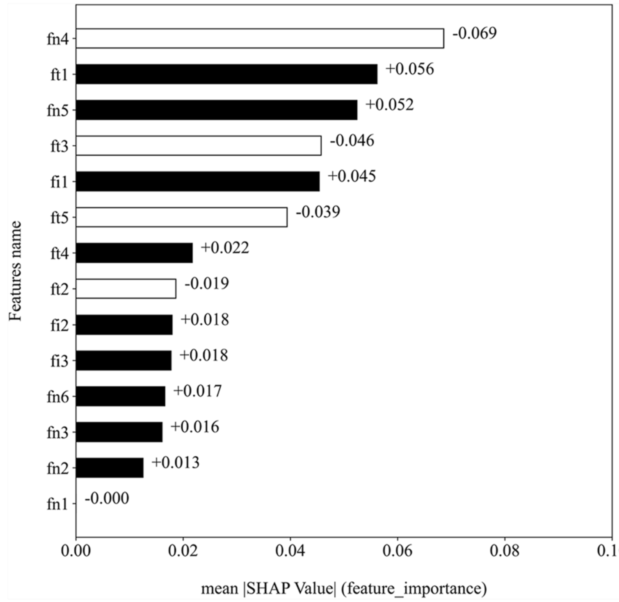

Fig. 4 Feature importance ranked by mean absolute SHAP values (positive and negative contributions). Note: the figures next to the bar chart are the mean of absolute SHAP values per component across our data set and are sorted by decreasing importance. The different colors highlight the correlations between each component and the predicted outcome. The black color means a feature is positively correlated with the predicted outcome, and the white color means a negative correlation

This suggests that while text quality has improved, it is not uniformly leveraged by all models, possibly due to investors placing less emphasis on textual changes post-regulation or some models failing to fully leverage the improved textual data. Additionally, the mixed results may point to underlying differences in how models process and interpret textual information.

Regarding the image information (panel C), despite better image quality post-regulation, predictive accuracy declines for most models except GB-DT. This finding aligns with Hypothesis 1a, suggesting that visual signals are less critical in financial decision-making, even when improved.

Panel D presents the prediction results for bundled signals in a portfolio. All models show significant improvements post-regulation. BP-NN (1) achieves a 14% increase in accuracy from 0.62 to 0.76. This observed enhancement can be largely attributed to the overall improvement in transparency and credibility of information disclosure in the crowdfunding domain (as shown in Table 4), coupled with clearer regulatory frameworks, helps investors process diverse information more effectively.

In conclusion, the 2016 *Interim Measures* have significantly improved the information environment, enhancing signal clarity and reducing information asymmetry. Quantitative data remains the most valuable signal, with its predictive power strengthened post-regulation, as investors prioritize structured, reliable information. Textual signals, although improved, still play a supplementary role due to limited investor focus. Visual signals, while enhanced, continue to have minimal impact on prediction accuracy. For entrepreneurs, these findings highlight the importance of clear, credible quantitative disclosures to attract investors. While high-quality textual and visual signals support campaigns, they are less influential compared to robust quantitative information in signaling project quality and securing funding.

6.5.3 *Placebo tests*

To further validate our findings and isolate the impact of the 2016 *Interim Measures*, we conducted placebo tests using randomly selected dates as alternative benchmarks. This approach is consistent with the concept of placebo tests in econometrics, which are employed to verify the robustness of causal inferences (Imbens & Wooldridge, 2009). These tests help determine whether the observed improvements in predictive performance can be specifically attributed to the regulatory changes or if they might be due to other factors, such as the natural learning curve of investors or random fluctuations in the data (Block et al., 2018; Vismara, 2018). We selected random dates, Apr 2014, Aug 2015, and June 2016, as alternative benchmarks to compare predictive performance across different time periods.

Table 8 Impact of policy on difference between successful and failed campaigns

	Before regulation			After regulation		
Variables	Fail (34)	Success (61)	Mean-Diff	Fail (18)	Success (31)	Mean-Diff
Panel A: numerical						
Target	0.064	0.094	−0.031	−0.252	−0.110	−0.142
Startlimit	−0.111	0.080	−0.191	−0.103	0.025	−0.128
Valuation	0.265	−0.110	0.376	−0.025	−0.059	0.0340
Sharetrans	−0.233	0.206	−0.439*	−0.183	−0.044	−0.139
Firstholder	−0.438	−0.180	−0.258	0.069	0.794	−0.725**
CEO	0.265	0.377	−0.112	0.222	0.613	−0.391***
Internet	0.471	0.590	−0.120	0.444	0.742	−0.297**
Technique	0.118	0.066	0.052	0.111	0.032	0.079
Market	0.912	0.934	−0.023	0.889	0.968	−0.079
Member	0.294	−0.055	0.348**	−0.278	−0.053	−0.225
Entrepre	0.176	−0.093	0.269	0.174	−0.112	0.286
Industry	0.375	0.092	0.283	−0.355	−0.386	0.031
Bachelor	−0.144	−0.070	−0.074	−0.043	0.322	−0.365
Master	0.023	0.041	−0.018	0.283	−0.269	0.552*
Doctor	0.062	0.090	−0.028	−0.291	−0.076	−0.215
MBA	−0.066	0.145	−0.212	−0.367	0	−0.366
Abroad	−0.127	0.183	−0.310	−0.090	−0.170	0.080
Other	−0.116	0.142	−0.258	−0.318	0.0310	−0.349
Panel B: text						
Positive_t	0.971	0.902	0.069	0.833	0.806	0.027
Negative_t	0.029	0.066	−0.036	0.056	0.129	−0.073
Neutral_t	0	0.033	−0.033	0.111	0.065	0.047
Senti_title	0.079	0.065	0.014	0.115	−0.280	0.395
Length_title	−0.011	0.054	−0.065	0.393	−0.322	0.715***
Length_intro	0.147	0.208	−0.062	−0.519	−0.270	−0.249
Verb	0.090	0.115	−0.025	−0.528	−0.018	−0.511**
Noum	0.135	0.113	0.022	−0.362	−0.159	−0.203
Adjv	0.208	0.079	0.130	−0.137	−0.303	0.166
Readability	0.064	−0.231	0.294	0.275	0.225	0.051
Similar_fund	−0.116	−0.058	−0.058	0.106	0.180	−0.075
Similar_class	−0.226	0.227	−0.453*	−0.268	−0.042	−0.226
Similar_all	0.088	−0.014	0.102	−0.132	0.008	−0.140
Panel C: image						
Clarity	−0.144	−0.083	−0.061	0.683	−0.075	0.758**
Youth	0.012	−0.244	0.256	0.485	0.185	0.300*
Beauty	−0.179	−0.218	0.039	0.322	0.438	−0.116
Male	0.061	−0.085	0.146	−0.261	0.252	−0.512*
Happy	0.009	0.128	−0.119	−0.049	−0.232	0.183
Fear	−0.083	0.113	−0.197	−0.083	−0.083	0
Neutral_i	−0.051	−0.387	0.336**	0.334	0.623	−0.290
NumPho	−0.123	−0.257	0.133	0.478	0.362	0.115

The results presented in Table 10 provide evidence supporting the causal inference of the policy effect. By examining the predictive performance both before and after the August 2016 policy implementation and comparing these with randomly selected alternative dates, a distinct pattern emerges. Specifically, following the

Table 9 Impact of policy by signal types and combined portfolio

Classifier/data			Acc	Prec	Rec	F1
Panel A: numerical						
Base learner	DT	Before	0.55	0.58	0.55	0.56
		After	0.65	0.64	0.65	0.65
	BP-NN (1)	Before	0.52	0.50	0.52	0.51
		After	0.67	0.66	0.67	0.65
	BP-NN (2)	Before	0.52	0.50	0.52	0.51
		After	0.69	0.68	0.69	0.67
Ensemble learning	RF-DT	Before	0.45	0.39	0.45	0.41
		After	0.67	0.66	0.67	0.63
	GB-DT	Before	0.59	0.38	0.59	0.46
		After	0.69	0.68	0.69	0.67
	GB-BP (1)	Before	0.55	0.53	0.55	0.53
		After	0.67	0.66	0.67	0.63
	GB-BP (2)	Before	0.52	0.50	0.52	0.51
		After	0.67	0.66	0.67	0.63
Panel B: text						
Base learner	DT	Before	0.52	0.52	0.52	0.52
		After	0.57	0.60	0.57	0.58
	BP-NN (1)	Before	0.59	0.51	0.59	0.51
		After	0.57	0.59	0.57	0.58
	BP-NN (2)	Before	0.52	0.48	0.52	0.49
		After	0.61	0.61	0.61	0.61
Ensemble learning	RF-DT	Before	0.52	0.48	0.52	0.49
		After	0.57	0.56	0.57	0.56
	GB-DT	Before	0.48	0.45	0.48	0.46
		After	0.49	0.50	0.49	0.49
	GB-BP (1)	Before	0.66	0.78	0.66	0.55
		After	0.51	0.52	0.51	0.51
	GB-BP (2)	Before	0.66	0.78	0.66	0.55
		After	0.55	0.56	0.55	0.56
Panel C: image						
Base learner	DT	Before	0.69	0.68	0.69	0.68
		After	0.57	0.58	0.57	0.57
	BP-NN (1)	Before	0.62	0.58	0.62	0.53
		After	0.49	0.40	0.49	0.44
	BP-NN (2)	Before	0.69	0.71	0.69	0.64
		After	0.51	0.48	0.51	0.49
Ensemble learning	RF-DT	Before	0.72	0.72	0.72	0.72
		After	0.69	0.68	0.69	0.67
	GB-DT	Before	0.62	0.63	0.62	0.62
		After	0.69	0.70	0.69	0.65
	GB-BP (1)	Before	0.66	0.66	0.66	0.59
		After	0.53	0.50	0.53	0.51
	GB-BP (2)	Before	0.62	0.58	0.62	0.53
		After	0.49	0.40	0.49	0.44
Panel D: signal portfolio						

Table 9 (continued)

Classifier/data			Acc	Prec	Rec	F1
Base learner	DT	Before	0.48	0.47	0.48	0.48
		After	0.59	0.57	0.59	0.58
	BP-NN (1)	Before	0.62	0.60	0.62	0.61
		After	0.76	0.75	0.76	0.75
	BP-NN (2)	Before	0.62	0.60	0.62	0.61
		After	0.73	0.73	0.73	0.73
Ensemble learning	RF-DT	Before	0.52	0.36	0.52	0.42
		After	0.71	0.75	0.71	0.67
	GB-DT	Before	0.48	0.45	0.48	0.46
		After	0.59	0.58	0.59	0.59
	GB-BP (1)	Before	0.59	0.57	0.59	0.58
		After	0.67	0.68	0.67	0.68
	GB-BP (2)	Before	0.55	0.54	0.55	0.55
		After	0.63	0.64	0.63	0.64

Note: *DT* decision tree, *BP-NN* backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting

August 2016 policy change, there is a general increase in accuracy scores across all models and signal types. Notable improvements are observed in models such as DT, BP-NN (1), and BP-NN (2), which demonstrate substantial gains across panels A (numerical) and C (image). In panel B (text), the DT and BP-NN (2) models also show notable improvements.

In contrast, the results for the placebo dates, April 2014, August 2015, and June 2016, do not show consistent improvements. For instance, in April 2014, accuracy scores for DT, BP-NN (1), and BP-NN (2) dropped significantly in both panel B (text) and panel C (image). In panel A (numerical), BP-NN (1) and BP-NN (2) also experience declines, while DT remain stable. Similarly, June 2016 sees a general decrease in accuracy across all panels, with only marginal gains for BP-NN (1) and BP-NN (2) in panel B (text). In August 2015, accuracy scores for all three models decline in panel B (text), while results in panels A (numerical) and C (image) are mixed, showing notable decreases alongside some slight increases.

This pattern indicates that the substantial improvements in predictive accuracy are closely linked to the August 2016 policy implementation, as similar gains are not widely observed at other randomly selected dates. Therefore, enhanced decision-making efficiency and reduced noise in equity crowdfunding can be attributed to the 2016 Interim Measures, rather than to random factors or the natural progression of the market. The policy appears to have played a crucial role in improving the information environment in China's equity crowdfunding market.

7 Robustness tests

To address concerns about the small sample size, we conduct two robustness checks using the tests described in Section 5.2.

7.1 Reduced dimensions

First, we reduce the feature dimensions to focus on the most critical variables, aiming to minimize complexity and prevent overfitting. Specifically, we previously reported deriving 6 components from the initial 18 quantitative variables, 5 components from the initial 13 text variables, and 3 components from the initial 8 image variables (Table 7). Table 11 here presents the prediction accuracy of these 14 components when used with the same model to predict equity crowdfunding success. The analysis shows that incorporating these PCA components does not significantly alter our overall findings, indicating that the

Table 10 Classification results of bundled signals over different sample periods

Classifier/data			Acc	Acc	Acc	Acc
Panel A: numerical			08.2016	04.2014	08.2015	06.2016
Base learner	DT	Before	0.55	0.71	0.58	0.69
		After	0.65	0.71	0.67	0.66
	BP-NN (1)	Before	0.52	0.75	0.67	0.77
		After	0.67	0.59	0.66	0.68
	BP-NN (2)	Before	0.52	0.67	0.67	0.77
		After	0.69	0.62	0.67	0.64
Panel B: text						
Base learner	DT	Before	0.52	0.67	0.67	0.62
		After	0.57	0.55	0.59	0.58
	BP-NN (1)	Before	0.59	0.58	0.75	0.54
		After	0.57	0.52	0.52	0.56
	BP-NN (2)	Before	0.52	0.54	0.75	0.46
		After	0.61	0.50	0.52	0.51
Panel C: image						
Base learner	DT	Before	0.48	0.75	0.50	0.81
		After	0.59	0.50	0.55	0.47
	BP-NN (1)	Before	0.62	0.71	0.83	0.77
		After	0.76	0.55	0.57	0.53
	BP-NN (2)	Before	0.62	0.67	0.75	0.77
		After	0.73	0.55	0.56	0.53

Note: 08.2016 (August 2016) is the policy issue date of *Interim Measures*. We randomly select April 2014 (04.2014), Aug 2015 (08.2015), and June 2016 (06.2016) as the cutting point of our sample to test whether there is a significant improvement of Acc. value after the selected dates

model's performance remains consistent despite the dimensionality reduction.

7.2 Alternative sample

Second, we enhance our analysis by incorporating data from an additional platform to further validate our findings. JD Finance and Dreammove are two platforms that scholars often utilize when studying equity crowdfunding in China (Chen & Ma, 2023; Zhao et al., 2021).[6] Dreammove has accumulated a total of 182 financing projects, with a cumulative financing amount of 269.09 million RMB and a total of 43,762 certified investors since its establishment. In contrast, JD Finance has been involved in 103 projects, with financing exceeding 1267.54 million RMB, and it boasts the participation of 80,131 individuals in equity investments.[7] To assess the robustness of our results, we conduct a similar analysis using the projects from JD Finance based on our model. The details of this analysis can be found in Table 15 in the Appendix, which reports the mean differences in common features between these two platforms.

To assess the robustness of our findings, we employ two additional samples for analysis. First, we apply our models to conduct the same analysis on a sample of 93 projects from JD Finance. Second, we combine the samples from Dreammove and JD Finance, totaling 237 projects, based on common features. The results of these analyses are presented in panel A and panel B of Table 12, respectively. We cannot report the results for image information when analyzing the sample with JD Finance projects, because JD Finance is no longer publicly accessible. In general, our findings indicate

[6] These two platforms exhibit significant differences in their operational focus. JD Finance operates as an investor-led equity crowdfunding platform with a particular emphasis on the influence of lead institutional investors on project success. On the other hand, Dreammove does not feature well-known lead investors and places a greater emphasis on small and microcrow investors. The scale of projects on Dreammove is comparatively smaller than those on JD Finance.

[7] Our observations are less than 182 and 103 due to the missing data of some key variables.

that both quantitative and textual information hold predictive power for equity crowdfunding project outcomes. Quantitative information is more effectively utilized by the DT model, while BP-NN performs better with textual and bundled information. These findings align with our primary results.

Table 11 Robustness check: classification results of PCA

Classifier/data		Acc	Prec	Rec	F1
Numerical	DT	0.63	0.71	0.63	0.63
	BP-NN (1)	0.63	0.62	0.63	0.62
	BP-NN (2)	0.63	0.62	0.63	0.62
Text	DT	0.58	0.59	0.58	0.59
	BP-NN (1)	0.65	0.42	0.65	0.51
	BP-NN (2)	0.65	0.42	0.65	0.51
Image	DT	0.40	0.41	0.40	0.40
	BP-NN (1)	0.65	0.42	0.65	0.51
	BP-NN (2)	0.65	0.42	0.65	0.51
Bundled	DT	0.60	0.59	0.60	0.59
	BP-NN (1)	0.72	0.71	0.72	0.72
	BP-NN (2)	0.74	0.75	0.74	0.75

Note: *DT* decision tree, *BP-NN* backpropagation neural networks, *BP-NN (1)* single-layer BP-NN, *BP-NN (2)* two-layer BP-NN, *RF* random forest, *GB* gradient boosting

8 Conclusion

Our study explores the interplay between signaling and perception in shaping the investment decisions of crowd investors in crowdfunding projects. While classical signaling theory typically focuses on how different types of signals influence investment decisions from the senders' perspective (Bi et al., 2017; Calic & Shevchenko, 2020), our findings emphasize the crucial role of signal perception from the receivers' standpoint in determining crowdfunding success. To replicate human thought processes, we use DT and BP-NN models as base learners and further enhance their predictive accuracy with ensemble algorithms.

A key insight from our research is the superior effectiveness of quantitative information over qualitative data, such as text and images, in conveying valuable insights to investors. The structured nature of quantitative data demands less cognitive effort compared to the complexity and ambiguity of qualitative information. When processed together, quantitative and qualitative signals increase complexity, necessitating higher analytical effort. This suggests that investors tend to allocate more attention to quantitative data, which enhances its predictive power in equity crowdfunding. From an institutional perspective, we find that regulatory measures significantly impact the effectiveness of signals in crowdfunding. Enhanced regulations focused on information disclosure help mitigate noise in the crowdfunding environment, thereby improving the quality of disclosed information. This leads to better information processing and more informed investment decisions. Our findings demonstrate that regulations promoting transparency and standardization in disclosures contribute to a more stable and trustworthy crowdfunding market.

Our research builds on Johan and Zhang (2020) by assessing the effectiveness of three distinct types of signals—quantitative, text, and image—within the fundraising context. We extend beyond the examination of qualitative business information alone by selecting data processing methods tailored to each information type's unique characteristics. This enables us to evaluate information processing effectiveness based on complexity and ambiguity. The study offers practical implications for entrepreneurs, investors, and policymakers. Entrepreneurs should prioritize providing clear and concise quantitative information to attract investors. Investors, in turn, should use appropriate models for each signal type to optimize decision-making. Policymakers can improve the crowdfunding ecosystem through regulations that enhance signal quality and reduce information asymmetry, fostering sustainable development in the sector.

However, our study has limitations related to sample size and the use of machine learning models in addressing endogeneity issues. Our sample size is relatively small, with 144 projects included in our main test due to data availability from Dreammove and the constraints imposed by the *Interim Measures*. To augment the sample size, we incorporate data from JD Finance for robustness check, adding 93 projects, though differences in data formats pose challenges. Future research could benefit from aggregating data across multiple platforms to expand the sample size (Cumming et al., 2019a; Rossi & Vismara, 2018). Additionally,

Table 12 Robustness check: classification results over different samples

Classifier/data		Acc	Prec	Rec	F1
Panel A: sample from JD Finance					
Numerical	DT	0.82	0.81	0.82	0.81
	BP-NN (1)	0.75	0.66	0.75	0.70
	BP-NN (2)	0.79	0.67	0.79	0.72
Text	DT	0.68	0.74	0.68	0.70
	BP-NN (1)	0.75	0.77	0.75	0.76
	BP-NN (2)	0.71	0.76	0.71	0.73
Bundled	DT	0.82	0.67	0.82	0.74
	BP-NN (1)	0.79	0.75	0.79	0.76
	BP-NN (2)	0.79	0.75	0.79	0.76
Panel B: an integrated sample from two platforms					
Numerical	DT	0.70	0.49	0.70	0.58
	BP-NN (1)	0.69	0.59	0.69	0.59
	BP-NN (2)	0.71	0.69	0.71	0.65
Text	DT	0.59	0.62	0.59	0.60
	BP-NN (1)	0.70	0.65	0.70	0.61
	BP-NN (2)	0.70	0.65	0.70	0.61
Bundled	DT	0.70	0.49	0.70	0.58
	BP-NN (1)	0.71	0.72	0.71	0.72
	BP-NN (2)	0.66	0.66	0.66	0.66

Note: the projects in JD Finance in our sample all achieved the target amount, and many exceeded the target amount. We code output variable as 1 if the actual amount exceeded target amount, otherwise coded as 0

standard machine learning estimators often struggle to identify causal relationships when dealing with endogenous regressors. In our study, we use machine learning methods with high-dimensional input features, which focus more on learning correlations to improve predictive performance rather than estimating specific causal parameters, as traditional econometric models do (Athey & Imbens, 2019). Although we attempt to mitigate endogeneity concerns by using the *Interim Measures* policy for identifying subgroups with different treatment effects and conducting placebo tests, endogeneity issues cannot be entirely resolved in our study. Future research could expand the sample size and employ methodologies suggested by Guo et al. (2022) to draw causal inferences about policy effects or investigate the impact of specific variables on our findings using explainable machine learning or traditional econometric models. Furthermore, we assume that investors uniformly perceive regulatory policies. However, it is crucial to acknowledge that investors' perceptions of regulatory policies may affect their views on risk assessment, compliance, and crowdfunding investment decisions. Future research could leverage machine learning techniques to identify patterns and sentiments related to regulatory policies and assess how these perceptions influence investment decisions.

Author contribution Jinjuan Yang: writing—original draft, software, methodology, data curation, and formal analysis. Jiayuan Xin: writing—review and editing, project administration, investigation, and conceptualization. Yan Zeng: writing—original draft, review and editing, project administration, supervision, and conceptualization. Pei Jose Liu: writing—review and editing, validation, visualization, and conceptualization.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability Data available on request from the authors.

Declarations

Competing interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix 1

Table 13 Platforms with fundraising more than 50 million in 2016

Platforms	Website	Status
JD Finance (京东东家)	https://dj.jd.com	No new project
Renrentou (人人投)	http://www.renrentou.com	Closed
Zhijinhui (智金汇)	http://www.zhijin.vc	Closed
Zhongtoubang (众投邦)	https://www.zhongtou8.cn	Closed
Zhongtoutiandi (众投天地)	http://www.coinvs.com	Closed
Zhongchouke (众筹客)	http://zhongchouke.com	No new project
36Kr (36氪)	https://www.36kr.com	Closed
Jingbeizhognchou(京北众筹)	http://www.jbzhongchou.com	Closed
Niutou (牛投网)	http://www.niutou.com	Closed
Aijiutou (爱就投)	http://www.i9tou.com	Transformation
The 5th avenue (第五创)	http://www.d5ct.com	Closed
Antsdaq (蚂蚁达客)	https://Antsdaq.com	No new project
Mic Finance (米筹金服)	https://www.mifund.com	Closed
Touwho (投壶网)	http://www.touwho.com	Closed
360taojin (360淘金)	https://wwww.360taojin.com	Closed
Qierong (企e融)	https://www.qierong.com	Closed
Dahuotou (大伙投)	https://www.dhtou.com	Closed
Choudao (筹道股权)	https://www.choudao.com	Closed
Dreammove (聚募网)	https://www.dreammove.cn	No new project
Kaoputou (靠谱投)	http://www.canyandata.com	Closed

Note: there are 200 platforms have emerged in the process of equity crowdfunding development in China. We surveyed the top 20 equity crowdfunding platforms by fundraising amount in 2016 to confirm whether their official websites are still alive. Table 13 shows the list of 20 equity crowdfunding platforms. The official websites of 15 platforms are not alive (closed). Among the 5 platforms that can be opened, Aijiutou has realized the transformation from an equity crowdfunding platform to a platform for entrepreneurial resources sharing; Zhongchouke cannot register and log in. We cannot register in JD Finance, Antsdaq and Dreammove, but old users of these platforms can still see the project information. Dreammove is one of the few platforms that the official websites can still open and carry out account balance withdrawals after the market elimination and government supervision shutdown, which passed the test of time

Table 14 Comparison of industry distribution

By operating sectors	Dreammove		32 platforms	
	Freq	Percent	Freq	Percent
Intellectual products manufacturing	8	5.56	115	11.49
Internet and modern IT service	36	25	296	29.57
Financial service	20	13.89	70	6.99
Business service	10	6.94	86	8.59
Culture, sports and leisure service	34	23.61	250	24.98
Travel service	9	6.25	31	3.10
Food and restaurant	20	13.89	75	7.49
Household convenience service	7	4.86	78	7.79
Total	144	100	1,001	100

Note: in 2016, there were 118 equity crowdfunding platforms in China, and 32 equity crowdfunding platforms of Z-park Crowdfunding Association were selected as our analysis samples (We choose the data for 2016 because in October 2016, the China Securities Regulatory Commission, together with 15 ministries and commissions, issued Circular No. 29 to strengthen regulations and restrictions on the development of equity crowdfunding. Due to regulatory reasons and the immaturity of equity crowdfunding in China, the number of equity crowdfunding platforms began to decrease after 2016, and the research reports on equity crowdfunding industry in China are mostly based on the data in 2016. After 2017, it is difficult to obtain the overall data of the equity crowdfunding industry). By the end of 2016, the 32 platforms have a cumulative number of 1001 projects with amount of 7.154 billion RMB, and 26,100 investors involved (Data source: 2017 The development of internet crowdfunding in China from *Z-park Crowdfunding Association.*). Table 13 shows the comparison of industry distribution between Dreammove and 32 platforms. Internet and modern IT services have the most projects, followed by culture, sports, and leisure services. The industry distribution of the projects in Dreammove is consistent with that in 32 platform

Table 15 Mean difference of common features between two platforms

Variables	Mean (JD Finance)	Mean (Dreammove)	Mean-Diff
Success	0.817	0.641	0.176***
Target	990.4	122.5	867.853***
Valuation	4453	2432	2020.486**
Member	4.129	4.965	−0.836***
Entrepre	0.334	0.217	0.117***
Industry	0.471	0.783	−0.312***
MBA	0.091	0.026	0.065***
Abroad	0.113	0.050	0.063***
Other	0.215	0.944	−0.729***
Positive_t	0.946	0.887	0.059
Negative_t	0.022	0.070	−0.049*
Neutral_t	0.022	0.042	−0.021
Senti_title	1.914	1.599	0.315***
Length_title	15.57	9.634	5.936***
Length_intro	91.44	63.86	27.582***
Readability	1967	720.7	1246.663***
Similar_fund	0.019	0.264	−0.246***

Note: 64.1% of the projects in Dreammove have achieved the target amount, and the actual financing amount is greater than the target amount for 81.7% projects in JD Finance. The project size of Dreammove, such as the target amount (122.5 vs. 990.4) and the project valuation (2432 vs. 4453), are smaller than projects in JD Finance. The team size of Dreammove is significantly larger than that of JD Finance (4.965 vs. 4.129). Compared with JD Finance, the entrepreneurs of Dreammove have rich industry experience, but insufficient entrepreneurial experience, fewer MBA degrees and overseas experience, and a high proportion of figures used in project presentation. As for the project titles, Dreammove is shorter than that of JD Finance, and the sentiment is not as positive as JD Finance. In terms of the project introduction, the sentiment of Dreammove is not as high-pitched as JD Finance, but the difference between the two platforms is not significant. Compared with JD Finance, the project introduction of Dreammove is shorter and less readable, but the similarity to successful projects

Appendix 2

Algorithmic derivation formulas for machine learning models.

1. General process for machine learning

Supposing that the training dataset is $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, where x_i is input variable, and $y_i = \{-1,1\}$ is output variable.

Machine learning involves two processes: learning and prediction. In the learning process, we obtain a decision function $Y = \widehat{f}(X)$ by learning from the given training dataset. In the prediction process, the prediction system gives the corresponding output y_{N+1} by the model $Y_{N+1} = \widehat{f}(X_{N+1})$ for the input x_{N+1}.

Learning algorithm selects the best model in training dataset by minimizing the difference between model output $\widehat{f}(x_i)$ and sample output y_i. We use loss function $L(Y, f(x))$ to measure prediction error. We select the optimal model based on the criterion of minimizing the expectation of loss function. However, the joint distribution of input and output variables is unknown, we cannot calculate the expectation of loss function. Therefore, empirical risk $R_{\mathrm{emp}}(f) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i))$ is used instead of expected loss. In all, to find the optimal model is to solve the optimization problem for the target function $\min_{f \epsilon \Gamma} \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i))$.

2. Decision tree (DT)

Expression of DT. DT is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules, and each leaf node represents the outcome, all data being in root node before splitting (Breiman et al., 2017). Supposing that we have classified the feature space into J regions $R_1, R_2, \cdots, R_J$, there is a specified output value c_j in each region J, the classification tree model can be expressed as follows:

$$f(x) = \sum_{j=1}^{J} c_j I(x \epsilon R_j)$$

$$c_j = arg\max_{y_i \epsilon Y} \sum_{x_i \epsilon R_j} I(\widehat{f}(x_i) = y_i)$$

where x_i($\mathrm{i} = 1, 2, 3, \cdots N$) is the input feature vector of training set, and $y_i \in Y = \{0,1\}$ is the prediction label of training set. $I(\bullet)$ is an indicator function that equals one if x is in R_j and zero otherwise. The optimal c_j is calculated by the rule of "majority vote," which means the mostly commonly occurring class in the classification process. It returns the actual label of y_i when the number of correct predictions ($\widehat{f}(x_i) = y_i$) is maximized in each region R_j

Generation of DT. We use Gini index to generate decision tree. For binary classification, if the probability of sample belonging to the first class is p, then the Gini index of the probability distribution is

$$\mathrm{Gini}(p) = 2p(1 - p)$$

We divide sample set D into D_1 and D_2 based on whether the feature A equals a:

$$\boldsymbol{D_1 = \{(x, y) \epsilon D | A(x) = a\}, D_2 = D - D_1}$$

Condition on feature A, the Gini index of set D can be defined as follows:

$$\mathrm{Gini}(D, A) = \frac{|D_1|}{|D|}\mathrm{Gini}(D_1) + \frac{|D_2|}{|D|}\mathrm{Gini}(D_2)$$

Then, we select the feature with the smallest Gini index and its corresponding critical value as the optimal feature and splitting point.

Pruning of DT. If we take 0–1 loss function for example, the target function is as follows:

$$\min_{J} \sum_{j=1}^{|J|} \sum_{x_i \epsilon R_j} I\left(y_i \neq \widehat{f}(x_i)\right) + \lambda \bullet |J|$$

where J represents subtree, and it is contained in the biggest tree $J_{\max}$ ($J \subseteq J_{\max}$). The complexity of subtree J is the number of terminal nodes $|J|$. $\lambda \bullet |J|$ is the penalization term for the complexity of tree. Then, we minimize the sum of cost and complexity penalty terms by search for the optimal λ:

$$\lambda = \min(\lambda, g(j))$$

$$g(t) = \frac{C(j) - C(R_j)}{\left|R_j\right| - 1}$$

where $C(j)$ and $C(R_j)$ are the empirical function of single node j and root node j, respectively. $\left|R_j\right|$ is the number of leaf node in subtree R_j.

3. Backpropagation neural network (BP-NN)

BP-NN is developed by Rumelhart et al. (1986). The expression of BP-NN is as follows:

$$G(x) = \sum_{i=1}^{m} \alpha_i f(w_i{\prime}x + b_i)$$

where (w_i, b_i) are the weight and intercept parameters for neural unit i. $f(\bullet)$ is the activation function, and α_i is the parameter to connect hidden layer and output layer. m represents the number of neural units.

In order to find the optimal parameter W^* for W (parameter vector contains all parameters in the BP-NN), we need to minimize the cost function:

$$W^* = \underset{W}{\mathrm{argmin}} \frac{1}{n} \sum_{i=1}^{n} L(y_i, G(x_i; W))$$

Then, we use gradient descent (partial derivative for W by back propagation) to update W, until the error term $\frac{\partial L}{\partial(f(w_i{\prime}x+b_i))}$ equals 0:

$$W = W - \eta \frac{\partial L(y_i, G(x_i; W))}{\partial W}$$

4. Ensemble learning

We combine base learner to construct a strong learner, which called ensemble learning, such as RF and GB (Friedman, 2001). The binary classification algorithm of GB is as follows:

First, we use the prior probability to initialize the first weak base learner as.

$$\boldsymbol{F_0(x) = \log \frac{P(Y=1|x)}{1-P(Y=1|x)}}$$

where $P(Y = 1|x)$ is the sample ratio of $y = 1$ in training sample.

Then, for $m = 1,2,\cdots,M$, we do the following loop:

a. Calculating the pseudo residuals for base learner m for $i = 1,2,\cdots,N$

$$r_{m,i} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x)}\right]_{F(x)=F_{m-1}(x)} = y_i - \frac{1}{1+e^{-F(x_i)}}$$

b. Based on $r_{m,i}$ to fit a tree, get the leaf nodes area $R_{m,j}, j = 1,2,\cdots,J$ for base learner m.

c. For $j = 1,2,\cdots,J$, calculate

$$c_{m,j} = \frac{\sum_{x_i \epsilon R_{m,j}} r_{m,i}}{\sum_{x_i \epsilon R_{mj}} (y_i - r_{m,i})(1 - y_i + r_{m,i})}$$

d. Replace the strong learner as

$$F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J} c_{m,j} I(x_i \epsilon R_{m,j})$$

Finally, we get the expression for the strong learner:

$$F_M(x) = F_0(x) + \sum_{m=1}^{M} \sum_{j=1}^{J} c_{m,j} I(x_i \epsilon R_{m,j})$$

References

Ahlers, G. K. C., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice, 39*, 955–980. https://doi.org/10.1111/Etap.12157

Allison, T. H., Davis, B. C., Webb, J. W., & Short, J. C. (2017). Persuasion In crowdfunding: An elaboration likelihood model of crowdfunding performance. *Journal of Business*

Venturing, 32, 707–725. https://doi.org/10.1016/J.Jbusvent.2017.09.002
Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General, 136*, 569–576. https://doi.org/10.1037/0096-3445.136.4.569
Anglin, A. H., Short, J. C., Drover, W., Stevenson, R. M., Mckenny, A. F., & Allison, T. H. (2018). The power of positivity? The influence of positive psychological capital language on crowdfunding performance. *Journal of Business Venturing, 33*, 470–492. https://doi.org/10.1016/J.Jbusvent.2018.03.003
Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4*(40–79), 40. https://doi.org/10.1214/09-Ss054
Asay, H. S., Libby, R., & Rennekamp, K. (2018). Firm performance, reporting goals, and language choices in narrative disclosures. *Journal of Accounting and Economics, 65*, 380–398. https://doi.org/10.1016/J.Jacceco.2018.02.002
Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics, 11*, 685–725. https://doi.org/10.1146/Annurev-Economics-080217-053433
Bafera, J., & Kleinert, S. (2023). Signaling theory in entrepreneurship research: A systematic review and research agenda. *Entrepreneurship Theory and Practice, 47*, 2419–2464. https://doi.org/10.1177/10422587221138489
Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (But Not True) news headlines. *Journal of Experimental Psychology: General, 149*, 1608–1613. https://doi.org/10.1037/Xge0000729
Bapna, S. (2019). Complementarity of signals in early-stage equity investment decisions: Evidence from a randomized field experiment. *Management Science, 65*, 933–952. https://doi.org/10.1287/Mnsc.2017.2833
Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies, 21*, 785–818. https://doi.org/10.1093/Rfs/Hhm079
Bayoudh, K., Knani, R., Hamdaoui, F., & Mtibaa, A. (2022). A survey on deep multimodal learning for computer vision: Advances, trends, applications, and datasets. *Visual Computer, 38*, 2939–2970. https://doi.org/10.1007/S00371-021-02166-7
Bergh, D. D., Connelly, B. L., Ketchen, D. J., & Shannon, L. M. (2014). Signalling theory and equilibrium in strategic management research: An assessment and a research agenda. *Journal of Management Studies, 51*, 1334–1360. https://doi.org/10.1111/Joms.12097
Bi, S., Liu, Z., & Usman, K. (2017). The influence of online information on investing decisions of reward-based crowdfunding. *Journal of Business Research, 71*, 10–18. https://doi.org/10.1016/J.Jbusres.2016.10.001
Block, J., Hornuf, L., & Moritz, A. (2018). Which updates during an equity crowdfunding campaign increase crowd participation? *Small Business Economics, 50*, 3–27. https://doi.org/10.1007/S11187-017-9876-4
Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39*, 713–726. https://doi.org/10.1016/S0896-6273(03)00466-5
Breiman, L. (1996). Bagging predictors. *Machine Learning, 24*, 123–140. https://doi.org/10.1007/Bf00058655
Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). Classification and regression trees. *New York, Chapman and Hall/crc*. https://doi.org/10.1201/9781315139470
Brown, D. E., Corruble, V., & Pittard, C. L. (1993). A comparison of decision tree classifiers with backpropagation neural networks for multimodal classification problems. *Pattern Recognition, 26*, 953–961. https://doi.org/10.1016/0031-3203(93)90060-A
Busenitz, L. W., Fiet, J. O., & Moesel, D. D. (2005). Signaling in venture capitalist—new venture team funding decisions: Does it indicate long-term venture outcomes? *Entrepreneurship Theory and Practice, 29*, 1–12. https://doi.org/10.1111/J.1540-6520.2005.00066.X
Butticè, V., & Vismara, S. (2022). Inclusive digital finance: The industry of equity crowdfunding. *The Journal of Technology Transfer, 47*, 1224–1241. https://doi.org/10.1007/S10961-021-09875-0
Butticè, V., Collewaert, V., Stroe, S., Vanacker, T., Vismara, S., & Walthoff-Borm, X. (2022). Equity crowdfunders' human capital and signal set formation: Evidence from eye tracking. *Entrepreneurship Theory and Practice, 46*, 1317–1343. https://doi.org/10.1177/10422587211026860
Calic, G., & Mosakowski, E. (2016). Kicking off social entrepreneurship: How a sustainability orientation influences crowdfunding success. *Journal of Management Studies, 53*, 738–767. https://doi.org/10.1111/Joms.12201
Calic, G., & Shevchenko, A. (2020). How signal intensity of behavioral orientations affects crowdfunding performance: The role of entrepreneurial orientation in crowdfunding business ventures. *Journal of Business Research, 115*, 204–220. https://doi.org/10.1016/J.Jbusres.2020.04.060
Cappa, F., Pinelli, M., Maiolini, R., & Leone, M. I. (2020). "Pledge" me your ears! the role of narratives and narrator experience in explaining crowdfunding success. *Small Business Economics, 57*, 953–973. https://doi.org/10.1007/S11187-020-00334-Y
Chaiken, S., & Ledgerwood, A. (2012). A theory of heuristic and systematic information processing. *Handbook Of Theories Of Social Psychology, 1*, 246–266. https://doi.org/10.4135/9781446249215.n13
Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293–332. https://doi.org/10.1207/S1532690xci0804_2
Chen, X., & Ma, L. (2023). Lead investors' insider ownership and crowd investors' agency concerns in investor-led equity crowdfunding. *Pacific-Basin Finance Journal, 78*, 101978. https://doi.org/10.1016/J.Pacfin.2023.101978
Chen, S., Duckworth, K., & Chaiken, S. (1999). Motivated heuristic and systematic processing. *Psychological Inquiry, 10*, 44–49. https://doi.org/10.1207/S15327965pli1001_6
Chen, X.-P., Yao, X., & Kotha, S. (2009). Entrepreneur passion and preparedness in business plan presentations: A

persuasion analysis of venture capitalists' funding decisions. *Academy of Management Journal, 52*, 199–214. https://doi.org/10.5465/Amj.2009.36462018

Cohen, L., Malloy, C., & Nguyen, Q. (2020). Lazy prices. *The Journal of Finance, 75*, 1371–1415. https://doi.org/10.1111/Jofi.12885

Colombo, M. G., Meoli, M., & Vismara, S. (2019). Signaling in science-based Ipos: The combined effect of affiliation with prestigious universities, underwriters, and venture capitalists. *Journal of Business Venturing, 34*, 141–177. https://doi.org/10.1016/J.Jbusvent.2018.04.009

Colombo, M. G., Fisch, C., Momtaz, P. P., & Vismara, S. (2022). The ceo beauty premium: Founder ceo attractiveness and firm valuation in initial coin offerings. *Strategic Entrepreneurship Journal, 16*, 491–521. https://doi.org/10.1002/Sej.1417

Cook, D. O., & Mobbs, S. (2022). Ceo selection and executive appearance. *Journal Of Financial And Quantitative Analysis, 58*, 1582. https://doi.org/10.1017/S0022109022000461

Correia, S., Sousa, M., & Brandão, E. (2024). What do we know about the choices of entrepreneurs before the equity crowdfunding campaign? *Small Business Economics*. https://doi.org/10.1007/S11187-023-00868-X

Courtney, C., Dutta, S., & Li, Y. (2017). Resolving information asymmetry: Signaling, endorsement, and crowdfunding success. *Entrepreneurship Theory and Practice, 41*, 265–290. https://doi.org/10.1111/Etap.12267

Cumming, D. J., Pandes, J. A., & Robinson, M. J. (2015). The role of agents in private entrepreneurial finance. *Entrepreneurship Theory and Practice, 39*, 345–374. https://doi.org/10.1111/Etap.12043

Cumming, D. J., Johan, S. A., & Zhang, Y. (2019a). The role of due diligence in crowdfunding platforms. *Journal of Banking & Finance, 108*, 105661. https://doi.org/10.1016/J.Jbankfin.2019.105661

Cumming, D. J., Vanacker, T. R., & Zahra, S. A. (2019b). Equity crowdfunding and governance: Toward an integrative model and research agenda. *Academy of Management Perspectives*. https://doi.org/10.5465/Amp.2017.0208

Curram, S., & Mingers, J. (1994). Neural networks, decision tree induction and discriminant analysis: An empirical comparison. *Journal of the Operational Research Society, 45*, 440–450. https://doi.org/10.1057/Jors.1994.62

Dane, E., & Pratt, M. G. (2007). Exploring Intuition And Its Role In Managerial Decision Making. *The Academy of Management Review, 32*, 33–54. https://doi.org/10.5465/Amr.2007.23463682

Danilov, A., & Sliwka, D. (2017). Can contracts signal social norms? experimental evidence. *Management Science, 63*, 459–476. https://doi.org/10.1287/Mnsc.2015.2336

De Crescenzo, V., Ribeiro-Soriano, D. E., & Covin, J. G. (2020). Exploring the viability of equity crowdfunding as a fundraising instrument: A configurational analysis of contingency factors that lead to crowdfunding success and failure. *Journal of Business Research, 115*, 348–356. https://doi.org/10.1016/J.Jbusres.2019.09.051

Di Pietro, F., Grilli, L., & Masciarelli, F. (2023). Talking about a revolution? costly and costless signals and the role of innovativeness in equity crowdfunding. *Journal of Small Business Management, 61*, 831–862. https://doi.org/10.1080/00472778.2020.1816435

Ding, C., Kavuri, A. S., & Milne, A. (2021). Lessons from the rise and fall of chinese peer-to-peer lending. *Journal of Banking Regulation, 22*, 133–143. https://doi.org/10.1057/S41261-020-00132-2

Drover, W., Wood, M. S., & Corbett, A. C. (2018). Toward a cognitive view of signalling theory: Individual attention and signal set interpretation. *Journal of Management Studies, 55*, 209–231. https://doi.org/10.1111/Joms.12282

Edelman, L. F., Manolova, T. S., Brush, C. G., & Chow, C. M. (2021). Signal configurations: Exploring set-theoretic relationships in angel investing. *Journal of Business Venturing, 36*, 106086. https://doi.org/10.1016/J.Jbusvent.2020.106086

Eldridge, D., Nisar, T. M., & Torchia, M. (2021). What impact does equity crowdfunding have on sme innovation and growth? An empirical study. *Small Business Economics, 56*, 105–120. https://doi.org/10.1007/S11187-019-00210-4

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, mis, and related disciplines. *The Information Society, 20*, 325–344. https://doi.org/10.1080/01972240490507974

Erel, I., Stern, L. H., Tan, C., Weisbach, M. S., & Goldstein, I. (2021). Selecting directors using machine learning. *The Review of Financial Studies, 34*, 3226–3264. https://doi.org/10.1093/Rfs/Hhab050

Eriksson, T., Bigi, A., & Bonera, M. (2020). Think with me, or think for me? On the future role of artificial intelligence in marketing strategy formulation. *The Tqm Journal, 32*, 795–814. https://doi.org/10.1108/Tqm-12-2019-0303

Evans, J. S. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*, 378–395. https://doi.org/10.3758/Bf03193858

Evans, J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology, 59*, 255–278. https://doi.org/10.1146/Annurev.Psych.59.103006.093629

Evans, J. S. (2010). Intuition and reasoning: A dual-process perspective. *Psychological Inquiry, 21*, 313–326. https://doi.org/10.1080/1047840x.2010.521057

Ferran, C., & Watts, S. (2008). Videoconferencing in the field: A heuristic processing model. *Management Science, 54*, 1565–1578. https://doi.org/10.1287/Mnsc.1080.0879

Franke, N., Gruber, M., Harhoff, D., & Henkel, J. (2008). Venture capitalists' evaluations of start–up teams: Trade-offs, knock-out criteria, and the impact of vc experience. *Entrepreneurship Theory and Practice, 32*, 459–483. https://doi.org/10.1111/J.1540-6520.2008.00236.X

Franzoni, C., & Tenca, F. (2023). How crowdfunders are influenced by entrepreneurial passion: A dual information processing perspective. *Entrepreneurship Theory and Practice, 47*, 1760–1787. https://doi.org/10.1177/10422587221102107

Friedman, J. H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics, 29*, 1189–1232. https://doi.org/10.1214/Aos/1013203451

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review*. *Psychology, 62*, 451–482. https://doi.org/10.1146/Annurev-Psych-120709-145346

Griffin, R. J., Neuwirth, K., Dunwoody, S., & Giese, J. (2004). Information Sufficiency And Risk Communication. *Media Psychology, 6*, 23–61. https://doi.org/10.1207/S1532785xmep0601_2

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*, 2223–2273. https://doi.org/10.1093/Rfs/Hhaa009

Guo, F., Huang, Y., Wang, J., & Wang, X. (2022). The Informal Economy At Times Of Covid-19 Pandemic. *China Economic Review, 71*, 101722. https://doi.org/10.1016/J.Chieco.2021.101722

Haefner, N., Wincent, J., Parida, V., & Gassmann, O. (2021). Artificial intelligence and innovation management: A review, framework, and research agenda. *Technological Forecasting and Social Change, 162*, 120392. https://doi.org/10.1016/J.Techfore.2020.120392

Hagras, H. (2018). Toward human understandable, explainable Ai. *Computer, 51*, 28–36. https://doi.org/10.1109/Mc.2018.3620965

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation, 16*, 45–74. https://doi.org/10.1007/S12559-023-10179-8

Hastie, T., Tibshirani, R., & Friedman, J. 2009. The Elements of Statistical Learning: Data Mining, Inference, And Prediction. Springer. Isbn-13: 978-0387848570

Hirshleifer, D., & Teoh, S. H. (2003). Limited Attention, Information Disclosure, And Financial Reporting. *Journal of Accounting and Economics, 36*, 337–386. https://doi.org/10.1016/J.Jacceco.2003.10.002

Hoegen, A., Steininger, D. M., & Veit, D. (2018). How do investors decide? an interdisciplinary review of decision-making in crowdfunding. *Electronic Markets, 28*, 339–365. https://doi.org/10.1007/S12525-017-0269-Y

Huang, R. H. (2018). Online P2p lending and regulatory responses in china: Opportunities and challenges. *European Business Organization Law Review, 19*, 63–92. https://doi.org/10.1007/S40804-018-0100-Z

Huang, Z., Chiu, C. L., Mo, S., & Marjerison, R. (2018). The nature of crowdfunding in china: Initial evidence. *Asia Pacific Journal of Innovation and Entrepreneurship, 12*, 300–322. https://doi.org/10.1108/Apjie-08-2018-0046

Huang, S., Pickernell, D., Battisti, M., & Nguyen, T. (2021). Signalling entrepreneurs' credibility and project quality for crowdfunding success: Cases from the kickstarter and indiegogo environments. *Small Business Economics*. https://doi.org/10.1007/S11187-021-00477-6

Huang, W., Vismara, S., & Wei, X. (2022). Confidence and capital raising. *Journal of Corporate Finance, 77*, 101900. https://doi.org/10.1016/J.Jcorpfin.2021.101900

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*, 5–86. https://doi.org/10.1257/Jel.47.1.5

Johan, S., & Zhang, Y. (2020). Quality Revealing Versus Overstating In Equity Crowdfunding. *Journal of Corporate Finance, 65*, 101741. https://doi.org/10.1016/J.Jcorpfin.2020.101741

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences, 374*, 20150202. https://doi.org/10.1098/Rsta.2015.0202

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. *Handbook of the Fundamentals of Financial Decision Making*. https://doi.org/10.1142/9789814417358_0006

Kahneman, D. (2011). Thinking, fast and slow. *Stat Papers* 55, 915 (2014). https://doi.org/10.1007/s00362-013-0533-y

Kaminski, J. C., & Hopp, C. (2019). Predicting outcomes in crowdfunding campaigns with textual, visual, and linguistic signals. *Small Business Economics, 55*, 627–649. https://doi.org/10.1007/S11187-019-00218-W

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society b: Biological Sciences, 361*, 2109–2128. https://doi.org/10.1098/Rstb.2006.1934

Kaur, D., Uslu, S., Rittichier, K. J., & Durresi, A. (2023). Trustworthy artificial intelligence: A review. *Acm Computing Surveys, 55*, 1–38. https://doi.org/10.1145/3491209

Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. *Journal of Applied Research in Memory and Cognition, 4*, 164–168. https://doi.org/10.1016/J.Jarmac.2015.07.001

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., et al. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances In Methods And Practices In Psychological Science, 1*, 443–490. https://doi.org/10.1177/2515245918810225

Kleinert, S., & Vismara, S. (2023). Quality Signals In Equity Crowdfunding. In D. Cumming & B. Hammer (Eds.), *The Palgrave Encyclopedia Of Private Equity*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-38738-9_184-1

Kleinert, S., & Volkmann, C. (2019). Equity Crowdfunding And The Role Of Investor Discussion Boards. *Venture Capital, 21*, 327–352. https://doi.org/10.1080/13691066.2019.1569853

Kunz, M. M., Bretschneider, U., Erler, M., & Leimeister, J. M. (2017). An empirical investigation of signaling in reward-based crowdfunding. *Electronic Commerce Research, 17*, 425–461. https://doi.org/10.1007/S10660-016-9249-0

Laureiro-Martínez, D., & Brusoni, S. (2018). Cognitive flexibility and adaptive decision-making: Evidence from a laboratory study of expert decision makers. *Strategic Management Journal, 39*, 1031–1058. https://doi.org/10.1002/Smj.2774

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444. https://doi.org/10.1038/Nature14539

Li, K., Mai, F., Shen, R., & Yan, X. (2021). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies, 34*, 3265–3315. https://doi.org/10.1093/Rfs/Hhaa079

Lukkarinen, A., & Schwienbacher, A. (2023). Secondary market listings in equity crowdfunding: The missing link? *Research Policy, 52*, 104648. https://doi.org/10.1016/J.Respol.2022.104648

Lundberg, S. M., & Lee, S. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (NIPS'17). Curran Associates Inc., Red Hook, pp. 4768–4777. https://doi.org/10.48550/Arxiv.1705.07874

Mahmood, A., Luffarelli, J., & Mukesh, M. (2019). What's In A Logo? The Impact Of Complex Visual Cues In Equity Crowdfunding. *Journal of Business Venturing, 34*, 41–62. https://doi.org/10.1016/J.Jbusvent.2018.09.006

Meoli, M., & Vismara, S. (2022). Machine-learning forecasting of successful Icos. *Journal Of Economics And Business, 121*, 106071. https://doi.org/10.1016/J.Jeconbus.2022.106071

Mishler, E. G. (1995). Models of narrative analysis: A typology. *Journal of Narrative and Life History, 5*, 87–123. https://doi.org/10.1075/Jnlh.5.2.01mod

Mochkabadi, K., & Volkmann, C. K. (2018). Equity crowdfunding: A systematic review of the literature. *Small Business Economics, 54*, 75–118. https://doi.org/10.1007/S11187-018-0081-X

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing, 29*, 1–16. https://doi.org/10.1016/J.Jbusvent.2013.06.005

Monin, P., Noorderhaven, N., Vaara, E., & Kroon, D. (2013). Giving sense to and making sense of justice in postmerger integration. *Academy of Management Journal, 56*, 256–284. https://doi.org/10.5465/Amj.2010.0727

Monsell, S. (2003). Task Switching. *Trends in Cognitive Sciences, 7*, 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives, 31*, 87–106. https://doi.org/10.1257/Jep.31.2.87

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*, 1–4. https://doi.org/10.1207/S15326985ep3801_1

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 534–552. https://doi.org/10.1037/0278-7393.14.3.534

Phan, P., Wright, M., & Lee, S.-H. (2017). Of robots, artificial intelligence, and work. *Academy of Management Perspectives, 31*, 253–255. https://doi.org/10.5465/Amp.2017.0199

Pineau, J., Vincent-Lamarre, P., Sinha, K., Larivière, V., Beygelzimer, A., D'alché-Buc, F., Fox, E., & Larochelle, H. (2021). Improving reproducibility in machine learning research (A Report From The Neurips 2019 Reproducibility Program). *Journal Of Machine Learning Research, 22*, 1–20. https://doi.org/10.48550/Arxiv.2003.12206

Piva, E., & Rossi-Lamastra, C. (2018). Human capital signals and entrepreneurs' success in equity crowdfunding. *Small Business Economics, 51*, 667–686. https://doi.org/10.1007/S11187-017-9950-Y

Plummer, L. A., Allison, T. H., & Connelly, B. L. (2016). Better together? Signaling interactions in new venture pursuit of initial external capital. *Academy of Management Journal, 59*, 1585–1604. https://doi.org/10.5465/Amj.2013.0100

Prokop, J., & Wang, D. (2022). Is there a gender gap in equity-based crowdfunding? *Small Business Economics, 59*, 1219–1244. https://doi.org/10.1007/S11187-021-00574-6

Rai, A. (2020). Explainable Ai: From black box to glass box. *Journal of the Academy of Marketing Science, 48*, 137–141. https://doi.org/10.1007/S11747-019-00710-5

Ralcheva, A., & Roosenboom, P. (2019). Forecasting success in equity crowdfunding. *Small Business Economics, 55*, 39–56. https://doi.org/10.1007/S11187-019-00144-X

Ranta, M., Ylinen, M., & Järvenpää, M. (2023). Machine learning in management accounting research: Literature review and pathways for the future. *European Accounting Review, 32*, 607–636. https://doi.org/10.1080/09638180.2022.2137221

Reuer, J. J., Tong, T. W., & Wu, C.-W. (2012). A signaling theory of acquisition premiums: Evidence from ipo targets. *Academy of Management Journal, 55*, 667–683. https://doi.org/10.5465/Amj.2010.0259

Rossi, A., & Vismara, S. (2018). What do crowdfunding platforms do? A comparison between investment-based platforms in europe. *Eurasian Business Review, 8*, 93–118. https://doi.org/10.1007/S40821-017-0092-6

Rossi, A., Vanacker, T., & Vismara, S. (2021). Equity crowdfunding: New evidence from US and UK markets. *Review of Corporate Finance, 1*, 407–453. https://doi.org/10.1561/114.00000009

Ruan, Q., Wang, Z., Zhou, Y., & Lv, D. (2020). A new investor sentiment indicator (Isi) based on artificial intelligence: A powerful return predictor in China. *Economic Modelling, 88*, 47–58. https://doi.org/10.1016/J.Econmod.2019.09.009

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by back-propagating errors. *Nature, 323*, 533–536. https://doi.org/10.1038/323533a0

Schneider, D., & Harknett, K. (2019). Consequences of routine work schedule instability for worker health and well-being. *American Sociological Review, 84*, 82–114. https://doi.org/10.1177/0003122418823184

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*, 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes, 53*, 252–266. https://doi.org/10.1016/0749-5978(92)90064-E

Simon, H. A. (1990). Bounded rationality. In: J. Eatwell, M. Milgate, & P. Newman (Eds.), *Utility and Probability*. The New Palgrave. Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-20568-4_5

Slovic, P. (1972). Psychological study of human judgment: Implications for investment decision making. *The Journal of Finance, 27*, 779–799. https://doi.org/10.1111/J.1540-6261.1972.Tb01311.X

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*, 355–374. https://doi.org/10.2307/1882010

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way To Prevent Neural Networks From Overfitting. *The Journal of Machine Learning Research, 15*, 1929–1958. Https://Jmlr.Org/Papers/V15/Srivastava14a.Html.

Steigenberger, N., & Wilhelm, H. (2018). Extending Signaling Theory To Rhetorical Signals: Evidence From Crowdfunding. *Organization Science, 29*, 529–546. https://doi.org/10.1287/Orsc.2017.1195

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological), 36*, 111–133. https://doi.org/10.1111/J.2517-6161.1974.Tb00994.X

Sugumaran, V., & Ramachandran, K. I. (2007). Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical Systems and Signal Processing, 21*, 2237–2247. https://doi.org/10.1016/J.Ymssp.2006.09.007

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure And abstraction. *Science, 331*(6022), 1279–1285. https://doi.org/10.1126/Science.1192788

Thiele, J. A., Faskowitz, J., Sporns, O., & Hilger, K. (2022). Multitask Brain Network Reconfiguration Is Inversely Associated With Human Intelligence. *Cerebral Cortex, 32*, 4172–4182. https://doi.org/10.1093/Cercor/Bhab473

Todd, P. M., & Gigerenzer, G. (2012). Ecological rationality: Intelligence in the world. *Oxford University Press*. https://doi.org/10.1093/Acprof:Oso/9780195315448.001.0001

Tran, V. T., Yang, B.-S., Oh, M.-S., & Tan, A. C. C. (2009). Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference. *Expert Systems with Applications, 36*, 1840–1849. https://doi.org/10.1016/J.Eswa.2007.12.010

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31*, 411–437. https://doi.org/10.1146/Annurev.Neuro.30.051606.094238

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. *Science, 185*, 1124–1131. https://doi.org/10.1126/Science.185.4157.1124

Utami, I. T., Sartono, B., & Sadik, K. (2014). Comparison of single and ensemble classifiers of support vector machine and classification tree. *Journal Of Mathematical Sciences And Applications, 2*, 17–20. https://doi.org/10.12691/Jmsa-2-2-1

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in Cognitive Sciences, 22*, 213–224. https://doi.org/10.1016/J.Tics.2018.01.004

Vergne, J., Wernicke, G., & Brenner, S. (2018). Signal incongruence and its consequences: A study of media disapproval and ceo overcompensation. *Organization Science, 29*, 796–817. https://doi.org/10.1287/Orsc.2018.1209

Vismara, S. (2016). Equity retention and social network theory in equity crowdfunding. *Small Business Economics, 46*, 579–590. https://doi.org/10.1007/S11187-016-9710-4

Vismara, S. (2018). Signaling To Overcome Inefficiencies In Crowdfunding Markets. In D. Cumming & L. Hornuf (Eds.), *The Economics Of Crowdfunding: Startups, Portals And Investor Behavior.* Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-66119-3_3

Vismara, S. (2019). Sustainability in equity crowdfunding. *Technological Forecasting and Social Change, 141*, 98–106. https://doi.org/10.1016/J.Techfore.2018.07.014

Wang, W., Mahmood, A., Sismeiro, C., & Vulkan, N. (2019). The evolution of equity crowdfunding: Insights from co-investments of angels and the crowd. *Research Policy, 48*, 103727. https://doi.org/10.1016/J.Respol.2019.01.003

Yan, W.-J., Wu, Q., Liang, J., Chen, Y.-H., & Fu, X. (2013). How Fast Are The Leaked Facial Expressions: The Duration Of Micro-Expressions. *Journal of Nonverbal Behavior, 37*, 217–230. https://doi.org/10.1007/S10919-013-0159-8

Yang, Y., Ruan, Q., Huang, S., Lan, T., & Wang, Y. (2021). Impact of the covid-19 outbreak on tourists' real-time on-site emotional experience in reopened tourism destinations. *Journal of Hospitality and Tourism Management, 48*, 390–394. https://doi.org/10.1016/J.Jhtm.2021.07.014

Zhao, Y., Xie, X., & Yang, L. (2021). Female entrepreneurs and equity crowdfunding: The consequential roles of lead investors and venture stages. *International Entrepreneurship and Management Journal, 17*, 1183–1211. https://doi.org/10.1007/S11365-020-00659-W

Zhi, R., Zhou, C., Li, T., Liu, S., & Jin, Y. (2021). Action unit analysis enhanced facial expression recognition by deep neural network evolution. *Neurocomputing, 425*, 135–148. https://doi.org/10.1016/J.Neucom.2020.03.036

Zhu, Y., Zhou, L., Xie, C., Wang, G.-J., & Nguyen, T. V. (2019). Forecasting smes' credit risk in supply chain finance with an enhanced hybrid ensemble machine learning approach. *International Journal of Production Economics, 211*, 22–33. https://doi.org/10.1016/J.Ijpe.2019.01.032