International Journal of E-Entrepreneurship and Innovation
Volume 14 • Issue 1 • January-December 2024

Innovative Reward-Based Crowdfunding Decision Model

Somboon Prasobpiboon
Chulalongkorn University, Thailand

Roongkiat Ratanabanchuen
Chulalongkorn University, Thailand

Achara Chandrachai
Chulalongkorn University, Thailand

Sipat Triukose
https://orcid.org/0000-0002-2826-752X
Chulalongkorn University, Thailand

ABSTRACT

Due to the risky nature of newly creative projects for entrepreneurs, reward-based crowdfunding is currently an alternative fundraising channel for those who need seed funding to finance the creation of their prototype. The objectives of this research are to explore the success factors, including entrepreneurial, project and campaign factors, in project fundraising under a reward-based crowdfunding platform. We propose to develop a model for predicting the success of crowdfunding projects by machine learning. The datasets have been retrospectively gathered from historical records of campaigns in the Kickstarter website. The study's findings show that the logistic regression and decision tree models, respectively, had accuracy rates of 88.2% and 88.8%. The highest accuracy percentage of 94.1% originates from new testing data that has been externally validated for the technology industry. The practical implication of our research is that entrepreneurs can apply the proposed prediction model to identify the most influential topical features embedded in campaigns.

KEYWORDS

Reward-Based Crowdfunding, Success Factors, Funding Contribution, Machine learning, Logistic Regression, Decision Tree

INTRODUCTION

Crowdfunding is a method of obtaining money for a project from many individuals in small funds, usually through an online platform that is a digital fundraising form for innovative projects and ventures, and it has become popular as an alternative to bridge the financing gap of new ventures over the past few years (Da Cruz, 2018). The provision of financial resources happens through these online platforms that act as digital intermediaries matching fundraisers and funders through crowdfunding campaigns managed by the platform (Belleflamme et al., 2015; Beier & Früh, 2020). New entrepreneurs are traditionally at greater risk and have a higher rate of failure in comparison with other businesses as there is uncertainty about the development of unproven products and services (Valanciene & Jegeleviciute, 2013).

DOI: 10.4018/IJEEI.349983

A growing number of creative people use internet crowdfunding websites to crowdfund their new projects (Belleflamme et al., 2014). Increasingly popular crowdfunding platforms like Kickstarter enable project creators to raise hundreds of millions of dollars annually. Every crowdfunding project has a well-defined goal to be achieved within a given time frame. If the funding goal is achieved within the deadline, then the project is successful. The fundraising model is all-or-nothing and, after the campaign's deadline, a campaign is only deemed successful if it has achieved its funding target. In this case, backers actually pay the money they have pledged, and the project idea is realized. In cases where the goal is not reached, the campaign has failed and no exchange of money takes place. The creator's time, money, and effort spent in the entire exercise of planning, designing, launching, and promoting the project are in vain.

As a result, project creators are eager to discover the likelihood of a project's success as soon as possible. Moreover, knowing the project's outcome in advance helps the users to work out future strategies. Creators whose campaigns have a low probability of success may want to improve their chances by providing updates and improving information related to entrepreneurial, project, and campaign factors. These features of projects are used by the project creator to evaluate the success probability of a project.

Four categories can be used to categorize crowdfunding: donation-based, reward-based, lending-based, and equity-based. In donation-based crowdfunding, donors contribute money with no intention of receiving anything in return (Guan, 2016), while in reward-based crowdfunding projects are proposed in exchange for nonmonetary prizes (Bi et al., 2017; Yu et al., 2018). Reward-based crowdfunding offers some nonmonetary benefits such as validation of the business idea, definition of the product or service (through customer feedback), product promotion (Da Cruz, 2018; Brown et al., 2017), innovation (Song et al., 2020), identification of internationalization opportunities (Ahsan & Musteen, 2021), or encouragement of sustainable development (Laurell et al., 2019). Instead of receiving nonfinancial benefits for their contribution, backers of lending-based crowdfunding receive interest income. In equity-based crowdfunding, the project creator provides a reward in the form of an equity stake so that backers can partake in the projects' profits. (Bannerman, 2013; Beaulieu et al., 2015). Reward-based crowdfunding differs from the equity-based crowdfunding model in the nature of the exchange. While individual investors in equity-based crowdfunding receive shares in exchange for their investment, funders in reward-based crowdfunding receive rewards or perks according to their contribution levels (Cavalcanti & Soetanto, 2022).

There are two categories of financing models in each crowdfunding platform: "all-or-nothing" and "keep-it-all." The first category enables project creators to get cash only when backers' contributions have reached the project's funding target within the campaign's time frame. If the contribution falls short of the declared funding target, nothing is given to the project creators. The second category, on the other hand, allows project founders to keep the entire financing contribution—even for unsuccessful ventures (Cumming et al., 2020).

Several items of research on crowdsourcing platforms have been released. Mollick (2014) offered an explanation of the factors that influence the success and failure of Kickstarter campaigns, provided various pieces of information on the factors that determine success, and examined the relationships between various campaign aspects and the results. Greenberg et al. (2013) proposed a success predictor for Kickstarter campaigns based solely on their static attributes, that is, attributes available at the launch of a campaign and achieved a 68% prediction accuracy. The objectives of this research were to explore the success factors, including entrepreneurial, project, and campaign factors, in project fundraising under reward-based crowdfunding platform. In this study, we concentrated on developing models for forecasting the success of Kickstarter campaigns using machine-learning approaches to ascertain the likelihood that a crowdfunding project would succeed. Therefore, the study addresses this gap by developing a model for predicting the success of reward-based crowdfunding projects by machine-learning techniques including logistic-regression and decision-tree algorithms that have been employed.

In Section 2 we discuss literature review and, in Section 3, we detail our data set, its key features, and the preprocessing we used. In Section 4, we demonstrate the several developed prediction models we created, outlining each model and demonstrating how well it performs. In Section 5, we discuss the findings of the research and, in Section 6, we conclude and supply practical implications.

LITERATURE REVIEW

Crowdfunding Model

Reward-based crowdfunding is a tool for stimulating innovation in entrepreneurs through the engagement of early adopters, considered as a crowdfunding form of open search, that is, actively seeking out ideas from outsiders (Leone et al., 2023), where backers are often early consumers in charge of revealing an evaluation of the project in terms of comments and improvements on the product or service before the market launch (Da Cruz, 2018) and actively participate in the innovation conversation proposing new features or uses for products and services. Reward-based crowdfunding platforms represent, in fact, an environment for community building, co-creation, crowdsourcing (Afuah & Tucci, 2012), market knowledge, open innovation, or the development of shared social identities (Brown et al., 2017; Beier & Früh, 2020). Brown et al. (2017) argued that platforms ensure a ready market for their new offerings; they boost brand image and gain support for brand-related causes. Paschen (2017) suggested that entrepreneurs resort to reward-based crowdfunding to identify not only the target market but also partners, distributors, and competitors, offering nonmonetary benefits in terms of validation of the business idea and product promotion (Brown et al., 2017; Gerber & Hui, 2013; Di Pietro et al., 2018) or pricing strategy (Sewaid et at., 2021). Besides developing innovation, reward-based campaigns can also help innovative teams develop their community around their project (Giones & Brem, 2019).

The present study focused on reward-based crowdfunding as it is the largest crowdfunding type in terms of the overall number of crowdfunding platforms as well as the funding amounts raised by being the fastest growing form of crowdfunding. Crowdfunding has just emerged as a viable option for business owners looking to finance innovative ventures, but it is still unclear how to make a project successful. The purpose of this study was to investigate the entrepreneurial, project, and campaign elements that contribute to project fundraising success on a platform for reward-based crowdfunding for start-up businesses.

Success Factors in Reward-Based Crowdfunding Projects

Several studies have been conducted on reward-based crowdfunding platforms with the focus on their features and forecasts of a successful project, according to our analysis of the literature on the components of successful projects in reward-based crowd financing projects. The success factors can be categorized into three elements, consisting of entrepreneurial, project, and campaign aspects. We also provide the rationale for our research hypotheses in each of the following independent variables.

Entrepreneurial Factor

Number of Created Projects. The ability of project creators to perform online engagement is critical for a successful crowdfunding campaign as it enhances the creator's profile and reputation (Chandna & Salimath, 2018). Entrepreneurs with prior crowdfunding experience, particularly successful experience, are more likely to be successful in their fundraising (Kim et al., 2017). This is because founders with greater expertise in crowdfunding have a much higher chance of establishing credibility, which leads to fundraising success. According to research by Daoyuan (2016), the number of founders who create projects has a positive and significant impact on the success of crowdfunding campaigns. This is because the founders' knowledge may increase because of the experience gained from creating projects, and this can also be viewed as experience with crowdfunding and entrepreneurship. The founder acquires the greatest abilities in how to present the project and discovers how to be successful in the fundraising process as a result of the learning effect that results from creating more projects.

Number of Project Backings of Project Initiator. It is important for project initiators to back a few projects before they start to create their own, for two reasons. First, previous project backings by the initiator send a signal to the crowd that a project initiator understands the dynamics of crowdfunding and believes in the idea of crowdfunding. Second, initiators can gain knowledge from their own experience as backers by supporting other projects, which enables them to recognize the elements that are important for inspiring confidence among potential backers (Kunz et al., 2016). According to Daoyuan (2016), research demonstrates that the number of backed projects, which is a measure of the reciprocity of project founders, has a favorable and significant impact on a project's performance. As a result, project initiators' involvement in other crowdfunding campaigns indicates their experience, which may have an impact on the behavior of supporters (Kunz et al., 2017).

Project Factor

Number of Images Used on Campaign Site. Supporters are more drawn to visuals since they make projects easier for them to understand. As a result, it is critical for creators to use images when interacting with the Kickstarter community and attracting backers. Therefore, more images are likewise favorably related to financing success (Zhou et al., 2016). The overuse of images, however, can cause a campaign page to become congested and ultimately have an unfavorable appeal or interfere with its readability (Kunz et al., 2017). In addition, the pictures included in the profile may be quite complex with many being actually of a textual context instead of being pure images and thus require more sophisticated techniques in processing in order to reveal the real effect of the picture (Daoyuan, 2016). According to this study, visually appealing images demonstrate an initiator's readiness to provide high-quality reward.

Having Video on Campaign Site. Video indicates quality because it communicates that the project initiator is confident in presenting the relevant product, which increases the likelihood that the project will succeed (Mollick, 2014). To help potential supporters understand the project and persuade them to support the campaign, project developers can aim to provide greater richness in material, such as videos (Daoyuan, 2016; Kunz et al., 2016).

Having a Project Website. An external website can expand the creators' channels for communicating with the crowd. An external website may be seen favorably, increasing the likelihood that a project will succeed (Kunz et al., 2017). Projects not featuring their own website or those on a social media platform have a substantially lower success probability (Müllerleile & Joenssen, 2015)

Number of Available Reward Levels. Founders can choose the number of reward levels that require the amount of money backers need to pledge to receive a reward. A higher number of different rewards increases the likelihood of project success (Kunz et al., 2016). It can further be viewed that the more different rewards a project offers, the more options a potential backer has to pick from, which ultimately results in more financial support (Kunz et. al., 2017).

Campaign Factor

Campaign Duration. A prolonged duration of fundraising can suggest an unreliable project narrative, which may lead to a decline in support (Frydrych et al., 2014). According to Mollick (2014), long campaign durations signal a lack of confidence in the project initiators to successfully raise money. Zhou et al. (2016) and Kunz et al. (2017) also demonstrated that a higher requirement of the funding goal and longer campaign duration are negatively associated with funding success. Therefore, a long duration may reduce the backer's confidence in the project initiator's ability to complete the project in the provided time and at a high quality.

Funding Goal. Realistic funding objectives should be established. Although the all-or-nothing funding model eliminates financial loss in case of project failure, the pledge goal remains a cue for project quality (Da Cruz, 2018). A lower funding target makes obtaining finance easier (Müllerleile & Joenssen, 2015) because a higher funding target suggests that the project's originator will have to put in more effort to justify the sought for funding (Frydrych et al., 2014). Additionally, Cordova et al. (2015) discovered that raising the amount requested lowers the likelihood that the entrepreneur will meet the funding target. In addition, Kim et al. (2017) showed that setting a high fundraising objective denotes the necessity for a sizable initial capital investment in a project, which is ultimately a signal of higher risk and might breed mistrust regarding the founder's ability to execute the project. Rational initiators with higher pledge goals typically have a more comprehensive understanding of the project compared to those with lower pledge goals (Wang et al., 2021). The pledge goal may serve as a cue for assessing project quality (Wang et al., 2021) and demonstrate entrepreneurs' estimations of the project's value. According to attribution theory, investors make causal inferences about project quality based on pledge goals from the investor's perspective (Banerjee & Bose, 2022; Li et al., 2020). Creators set higher pledge goals when they are confident in their quality judgment; otherwise, a higher pledge goal can hinder the success of the project (Da Cruz, 2018). Excessively high pledge goals may decrease investors' willingness to participate. Projects with high pledge goals are usually more difficult than those with lower pledge goals, and the resulting risks may prompt potential investors to face less risky investments in similar projects (Wang et al., 2021). Thus, the funding goal has a negative impact on a successful campaign (Kunz et al., 2017). However, there is a moderating effect of pledge goals on the influence of reward on investors' willingness to participate because a high pledge goal enhances investors' willingness to participate; an extremely high pledge goal presents negative effects and can cause risk perception, according to Wang et at. (2024).

Delivery Time of Reward. Delivery time is considered to be a crucial indicator of service quality in the context of online services. Different from traditional markets, crowdfunding project rewards are usually time-delayed, and the public is motivated to switch from traditional ones to crowdfunding markets mainly because crowdfunding products or services are usually novel or original (Banerjee & Bose, 2022; Luo et al., 2022; St. John et al., 2022). A product or service may be near to completion if delivery times are short. Additionally, it functions as a quality signal that lowers perceived risks and increases perceived trust, which influences funding behavior (Kunz et al., 2017). The project initiators convey their confidence and capacity to fulfill the rewards as promised and on time by setting delivery dates that are near to the end of a campaign. On the other hand, funders looking for new projects may view a lengthy delivery time as risky and view initiators as being underprepared to see their own project through to completion.

All success variables for reward-based crowdfunding projects and their relationship to financing success are summarized in Table 1.

Machine-Learning Techniques in Prediction Model

Machine learning is a branch of artificial intelligence and computer science that focuses on using data and algorithms to simulate how humans learn, gradually improving the accuracy of the system. Machine learning is an umbrella term for problem-solving techniques that use algorithms to carry out necessary tasks. The mathematical foundations of machine learning are provided by mathematical optimization methods. The supervised learning model used in this study separated the data into areas with linear boundaries. The white circles and the black circles are separated here by a linear barrier. Using inputs and desired outputs, supervised learning algorithms create a mathematical model of a set of data.

The data are a collection of training examples and are referred to as training data. A supervisory signal, also known as the desired output, is present in each training example along with one or more inputs. An optimal function allows the algorithm to correctly determine the output for inputs that are not a part of the training data. An algorithm is considered to have learned to do a task when, over time, the accuracy of its outputs or predictions increases. Machine-learning methods usually exhibit stronger performances in out-of-sample data predictions (Platt, 1999), by applying methodology to

Table 1. Overview of variables in reward-based crowdfunding

Description		Relationship to funding success
1 Entrepreneurial factor		
1.1	Number of created projects (EA1)	Positive
1.2	Number of project backings of the project initiator (EA2)	Positive
2 Project factor		
2.1	Number of images used on the campaign site (PA1)	Positive
2.2	Having a video on the campaign site (PA2)	Positive
2.3	Having a project website (PA3)	Positive
2.4	Number of available reward levels (PA3)	Positive
3 Campaign factor		
3.1	Campaign duration (CA1)	Negative
3.2	Funding goal (CA2)	Negative
3.3	Delivery time of reward (CA3)	Negative

reduce the overfitting problem with splitting the available data into separate sets, for the purpose of training and testing.

Machine-learning approaches are traditionally divided into three broad categories, which correspond to learning paradigms and are dependent on the nature of the "signal" or "feedback" available to the learning system. First, in supervised learning, the computer is presented with example inputs and their desired outputs, given by a "teacher," and the goal is to learn a general rule that maps inputs to outputs. Supervised learning can be applied with decision-tree and logistical-regression algorithms. Decision-tree learning uses a decision tree as a prediction model to go from observations about an item, represented in the branches, to conclusions about the item's target value, represented in the leaves. Classification trees are decision-tree models where the target variable can take a discrete range of values. In these tree structures, the leaves correspond to class labels and the branches to the feature conjunctions that result in those class labels. Decision trees, where the target variable can take continuous values, typically real numbers, are called regression trees. Regression trees are decision trees where the target variable can take continuous values, often real numbers. The logistic model is a binary logistic-regression model that models the probability of an event taking place by having the log-odds for the event be a linear combination of one or more independent variables. The alternate names are derived from the logit or logistic unit, which is the unit of measurement for the log-odds scale.

Second, in unsupervised learning, the learning algorithm is not provided with labels; instead, it must determine the structure of its data on its own. This method can be used as a means to an end (feature learning) or as a goal in and of itself (discovering hidden patterns in data). Last, reinforcement learning involves a computer program interacting with a dynamic environment in which it must accomplish a certain task (like driving a car or competing in a game). As it navigates its problem space, the program is provided with feedback that is analogous to rewards, which it tries to maximize (Bishop & Nasrabadi, 2006).

There have been some studies regarding the prediction of success in crowdfunding projects using machine-learning techniques. Research by Greenberg et al. (2013) analyzed the effect on project outcome of static features that are available from the launch of a project. The authors used a support vector machine (SVM) and decision trees, assisted with boosting, to build a classier model that achieves an accuracy of 68% and Etter et al. (2013) used Markov chain modeling, using the time variant features of amount of money pledged and the number of backers, combined with SVM

Table 2. Number of reward-based crowdfunding projects between January 2014 and February 2020

Industry	No. of reward-based crowdfunding projects					
	Successful projects		Unsuccessful projects		Total	
	No of project	%	No of project	%	No of project	%
Design	11,712	47%	13,229	53%	24,941	100%
Fashion	6,155	34%	12,077	66%	18,232	100%
Technology	6,788	25%	20,487	75%	27,275	100%
Total	24,655	35%	45,793	65%	70,448	100%

Note. Source: https://www.kickstarter.com.

modeling of the social features regarding the creators/backers of projects through their Twitter activity. The authors were able to predict the success of a project with 76% accuracy. Hussaing et al. (2018) applied random forest to be the classifier for the prediction of success in a crowdfunding model, providing an accuracy of 80.4%.

For this research, we used machine-learning techniques and applied a supervised learning approach with a classification model using logistical-regression and decision-tree algorithms to classify the successful and unsuccessful reward-based crowdfunding campaigns and to construct the prediction model to predict the success of crowdfunding initiatives. A confusion matrix is used to summarize the predicted outcome.

RESEARCH METHODOLOGY

Data Collection

Because Kickstarter is not only the industry leader in reward-based crowdfunding platforms, which contain relevant and accurate records, but also includes a sizable amount of public information to use in deep learning analysis, we gathered historical data from the Kickstarter website for this study. We also chose three industries (design, fashion, and technology) that are more closely tied to new entrepreneurs.

Thus, between January 2014 and February 2020, data were scraped from the Kickstarter website to create our data set. It consists of 70,448 projects, of which 24,655 were successful, making up 35% of all projects. Our data set can be divided into three industries, with 24,941 projects in the design industry making up 35% of all projects, 18,232 projects in the fashion industry making up 26% of all campaigns, and 27,275 projects in the technology industry making up 39% of all projects, as shown in Table 2.

Data Analysis

The fundamental statistical significance of the variables is observed in data analysis. Understanding the data dispersion and frequency distribution characteristics of observable variable data is the purpose of analysis. We concentrated on the study of mean, standard deviation, skewness, and kurtosis of data from observable variables. When skewness and kurtosis are analyzed, the data distribution is still not a normal distribution if the skewness or kurtosis values are more than three or ten, respectively (Kline, 2011). To normalize observed variables into normal distribution for model building, we therefore applied log 10 to some observable variable data that are not normally distributed, as shown in Table 3.

Table 3. Statistical data for collected data of observed variables with influence on backers' contributions in reward-based crowdfunding projects

Factors impact on the successful reward-based crowdfunding projects	Statistic Figures			
	Mean	S.D.	Skewness	Kurtosis
Entrepreneurial factor				
1 Experience to launch project (EA1) *	0.41	0.20	2.42	6.54
2 Experience to back other project (EA2) *	0.36	0.50	1.43	1.48
Project factor				
3 No. of images (PA1) *	0.79	058	-0.12	-1.26
4 Having a project video (PA2)	0.39	0.49	0.43	-1.81
5. Having a project website (PA3)	0.74	0.44	-1.12	-0.75
6. No. of available reward levels (PA4) *	0.85	0.27	-0.31	-0.14
Campaign factor				
7. Campaign duration (CA1)	34.01	11.42	0.95	1.23
8. Funding goal (CA2) *	3.92	0.76	-0.27	0.63
9. Delivery time of reward (CA3) *	1.86	0.38	-0.56	1.01

Note. *To normalize independent variables for model construction, log 10 of the raw data plus one was applied to the independent variables.

Correlation Analysis

To ascertain whether the observed variables were correlated, observable variable correlation analysis was performed. The Pearson correlation matrix was calculated. A statistical indicator of the strength of a linear link between two variables is the Pearson correlation coefficient. The correlation coefficient can range in value from -1 to +1 and can be either positive or negative. To analyze data, Berry and Feldman (1985) found that multicollinearity is not a problem for multicollinearity testing if no correlation exceeds 0.80. The Pearson correlation matrix was applied to represent correlation values between independent variables.

Our correlation analysis, as shown in Table 4, shows that there are no problems with multicollinearity for coefficient estimates because all correlation coefficients are between -0.27 and +0.54, which is not less than -0.8 and not greater than +0.8 according to Berry and Feldman (1985). As a result, we can use all observable variables to build the prediction model.

FINDINGS

Model Development

Using machine-learning techniques and a supervised learning approach by classification type, we developed a prediction model for a reward-based crowdfunding campaign in this study using both successful and unsuccessful reward-based crowdfunding project data to build the model. We built a prediction model using historical campaign results (label) and observable variable data from reward-based crowdfunding campaigns. We observed a total of 70,448 projects from the data collected for both successful and unsuccessful reward-based crowdfunding initiatives, which were divided into three industries (design, fashion, and technology). We gathered information from historical observable variable data and campaign results for each project.

In this study, we used the multistage sampling method with 600 samples per industry (external validation data) to divide the data for model testing by isolating the number of samples from all sample

Table 4. Pearson correlations

	1	2	3	4	5	6	7	8	9
1 EA1	1	0.45**	0.18**	0.14**	0.12**	0.11**	-0.10**	-0.27**	-0.05**
2 EA2		1	0.32**	0.29**	0.21**	0.31**	-0.08**	-0.11**	0.02**
3 PA1			1	0.54**	0.25**	0.58**	-0.01	0.11**	0.13**
4 PA2				1	0.22**	0.39**	-0.01	0.05**	-0.08**
5 PA3					1	0.26**	-0.03	0.06**	-0.02**
6 PA4						1	-0.01*	0.11**	0.08**
7 CA1							1	0.25**	0.05**
8 CA2								1	0.32**
9 CA3									1

Note. **significance at the 0.01 level; * significance at the 0.05 level.

data into a separated data set. To evaluate the external validity of the model created from training data, we then applied the sample random sampling method in each industry to obtain the number of successful crowdfunding projects at 300 samples per industry and another 300 unsuccessful projects per industry. Consequently, the total quantity of sample data remaining after extracting the data was utilized to evaluate the model. This left 68,648 samples, of which training data made up 70% of the remaining samples to construct the prediction model and testing data made up 30% of the remaining samples to validate the model. For model development in each industry, we selected training data using a random sampling method for each industry to have equal numbers of examples in both the successful and unsuccessful crowdfunding projects for each industry.

To distinguish between successful and unsuccessful reward-based crowdfunding campaigns, we built a prediction model for reward-based campaigns using machine-learning techniques with a supervised learning method. In this paper, we use two algorithms to construct a prediction model for reward-based crowdfunding campaigns that evaluated the likelihood of success. These two algorithms are logistic regression and decision tree.

Prediction Model Developed by Using Logistic-Regression Algorithm

A logistic-regression model was used, which enables a statistical evaluation of the influence of several variables on a dependent variable of a dichotomous type (Hosmer & Lemeshow, 2000; Szafraniec-Siluta et al., 2022). It is a widely used multidimensional method for modeling dichotomous results (Bennouna et al., 2019; Jain et al., 2020; Kuswanto et al., 2015; Zahi & Achchab, 2020). It is also suitable for models covering decision-making issues, which is why it is often used in statistical analyses appearing in economics and finance literature (Strzelecka et al., 2020). The maximum likelihood method was used to estimate the logistic-regression coefficients (Maddala, 2001).

For the construction of the model, the dependent variable (Y) is a categorical variable, thus if it has a value of 1, the campaign has been successful in raising the desired amount of money. However, if the dependent variable equals 0, the project has failed to raise the money. A successful project is defined as a project that can raise funds equal to or greater than its target funding. Independent variables were assigned to nine independent variables as shown in Table 5. By utilizing logistic-regression methods, we created a prediction model for reward-based crowdfunding using the training data as described. To build the prediction model, which is shown in Table 5, we used historical campaign results (label) and data from observable variables of reward-based crowdfunding campaigns. The estimated model in form of the logistic-regression model of the probability of successful crowdfunding campaigns is shown in Table 5.

Table 5. Prediction model from training data using logistic-regression algorithm

No.	Independent variable			Beta Coefficient
	Variable name	Variable abbreviation	Type of data	
	Constant value			1.812
Entrepreneurial factor				
1	Experience to launch project	EA1	Continuous	1.989
2	Experience to back other project	EA2	Continuous	3.507
Project factor				
3	Number of images	PA1	Continuous	1.396
4	Having a project video	PA2	Dichotomous	4.055
5	Having a project website	PA3	Dichotomous	0.519
6	Number of available reward levels	PA4	Continuous	3.639
Campaign factor				
7	Campaign duration	CA1	Continuous	-0.942
8	Funding goal	CA2	Continuous	-14.793
9	Delivery time of reward	CA3	Continuous	0.561

The probability of successful crowdfunding campaigns can be expressed by this formula:

$$Prob(Y) = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

$where : f(x)$

$$= 1.182 + 1.989(EA_1) + 3.507(EA_2) + 1.396\,(PA_1) + 4.055\,(PA_2) + 0.519(PA_3) + 3.639(PA_4) - 0.942\,(CA_1) - 14.793\,(CA_2) + 0.561(CA_3)$$

Experience launching a crowdfunding project (EA1) and experience backing another crowdfunding project (EA2) are independent variables in the entrepreneurial factor. The coefficient values for each variable are 1.989 and 3.507, respectively. Both independent variables have a positive relationship with the likelihood that a reward-based crowdfunding campaign will succeed. As a result, entrepreneurs who have previously launched their reward-based crowdfunding project are able to use the lessons learned from their experience to make improvements to their projects for upcoming crowdfunding campaigns and have a better chance of becoming successful. According to statistical analysis, entrepreneurs who have never run a crowdfunding campaign have a greater risk of failing at approximately 74% of those who have.

Entrepreneurs could also pay attention to other people's crowdfunding campaigns and contribute to them in order to gain greater exposure and see how they should organize and enhance their own future crowdfunding campaigns. Additionally, it demonstrates to backers what the project creators intend and how they contribute to other people's crowdfunding efforts. This encourages backers to support entrepreneurs' crowdfunding campaigns and makes them feel good about the people behind the projects.

The number of images (PA1), having a project video (PA2), having a project website (PA3), and the number of reward levels (PA4) are independent variables in the project factor. The coefficient values for each variable are 1.396, 4.055, 0.519 and 3.639, respectively. The likelihood of a reward-based crowdfunding campaign succeeding is positively connected with each of the four independent variables. Therefore, when launching a project, entrepreneurs should include comprehensive images of the project to allow the supporter to see the whole picture and comprehend the rewards that will be offered. This aids backers in deciding how much money to contribute to a crowdfunding campaign. According to statistical analysis, entrepreneurs who present 20 or more photos to show project details have a greater than 50% chance of succeeding in a reward-based crowdfunding campaign, whereas entrepreneurs who do not present any photos are up to 90% more likely to fail in their attempt to raise money through crowdfunding.

To make up for the fact that entrepreneurs do not get the chance to meet and speak with backers directly, they should create a video that explains the project details, including rewards, to assist backers to understand and feel more inclined to engage in crowdfunding campaigns.The statistical data analysis of project data suggests that up to 74% of all projects with video creation have a possibility of succeeding in their crowdfunding campaign when entrepreneurs record a video to communicate project details. Meanwhile, entrepreneurs have a 90% chance of failing in their crowdfunding campaigns if they do not provide video specifics about the project details.

To make it easier for project supporters to get additional project information, entrepreneurs can establish a project website that displays more project specifics. Additionally, it helps to establish trust and confidence and shows the aim of the project creators looking to raise money through crowdfunding. According to statistical analysis, entrepreneurs are 80% more likely to fail in their efforts to raise money through crowdfunding if they do not have a website for their campaigns. Additionally, entrepreneurs should provide a range of reward alternatives so that backers can pick prizes that are appropriate and meet their needs. This will help backers make decisions to support crowdfunding.

The campaign duration (CA1), the funding goal (CA2), and the delivery time of reward (CA3) are independent variables in the campaign factor. The campaign duration and funding goal variables have coefficient values of -0.942 and -14.793, respectively. The likelihood that a reward-based crowdfunding campaign will be successful is negatively associated with both independent variables. Therefore, entrepreneurs should appropriately establish the campaign duration. The total number of days for the campaign period should not be excessively long because this detracts from the project's appeal, delays backers' decisions to support the crowdfunding campaign, and may even cause them to forget to revisit the project. From historical data in this research, we find that the effective total number of days for the successful campaign period should be approximately 30. Additionally, the funding goal should be set at an appropriate level that corresponds to the project's size.

The crowdfunding campaign will receive less money from backers if the target funding goal is set too high. However, the variable for reward delivery time has a coefficient value of 0.561, demonstrating a slightly positive correlation with the likelihood of crowdfunding success. To provide project backers with confidence and trust that the reward can be delivered on time, the entrepreneur should decide the reward delivery time in accordance with the anticipated timing of raw material procurement and actual manufacturing because, if the operator defines too short a delivery time for rewards, it will make the funders lose trust that the project will be able to be completed until the reward is delivered on time and dissuade them from giving money to support that reward-based crowdfunding project.

This study employed the confusion matrix to analyze the logistic-regression algorithm's performance, and the findings are displayed in Table 6. It was discovered that the prediction model's accuracy value is 88.2%.

Model Development by Using Decision-Tree Algorithm

The decision-tree algorithm uses the attribute of the data in classification to learn by grouping the data into different categories. This technique creates a branch based on the feature's value to decide which characteristic will serve as the tree's root or node. The model was developed using the training data that had been collected. The decision-tree algorithm was used to develop prediction

Table 6. Efficiency results of prediction model from training data using logistic-regression algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of project	% of total project	No. of projects
Actual result	Unsuccessful campaign	20,431	42.5%	3,597	7.5%	24,028
	Successful campaign	2,100	4.4%	21,928	45.6%	24,028
	Total	22,531	46.9%	25,525	53.1%	48,056

Note. Remark: Accuracy rate is (20,431 + 21,928) / 48,056 = 88.2%.

Table 7. Efficiency results of prediction model from training data using decision-tree algorithm

Decision tree algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	20,859	43.4%	3,169	6.6%	24,028
	Successful campaign	2,212	4.6%	21,816	45.4%	24,028
	Total	23,071	48.0%	24,985	52.0%	48,056

Note. Remark: Accuracy rate is (20,859 + 21,816) / 48,056 = 88.8%.

model in reward-based crowdfunding to evaluate the chances of crowdfunding success by using the crowdfunding outcome data (label) and the observed variable data of closed-crowdfunding projects.

To summarize performance analysis results, the confusion matrix is utilized to analyze the performance of the decision-tree algorithm. According to the results, the model's accuracy value is 88.8%, as shown in Table 7.

Model Development for Each Industry by Using Logistic-Regression Algorithm

The researcher used the logistic-regression technique to construct a prediction model for crowdfunding for each business, which consists of the design, fashion, and technology industries. The logistic-regression method has the advantage of being simple to use, simple to interpret, and simple to explain. It can also explain the relationship between the variables in terms of the degree and the direction of the relationship, making it suitable for further commercial use. The development and testing of the logistic-regression model gave the model validity results that are similar to those obtained using the decision-tree method.

The researcher divided the data of reward-based crowdfunding projects that have already been launched in each industry into successful and unsuccessful campaigns. Such sample data includes details on both the observable variables and the results of the reward-based crowdfunding campaign that are utilized to build the model.

The researcher divided the data for testing the model by extracting the number of samples from the total sample data as a separated data set, using multistage sampling by dividing the data into each of the three industries and then determining the required samples for each industry at 600 samples per industry to be external validation data, which were used for each industry. Simple sample random sampling was used in each industry to obtain 300 successful projects in reward-based crowdfunding funding, along with another unsuccessful crowdfunding project of 300 samples to test the external validity of a model developed from training data. After that, the researcher divided the data by random

Table 8. Prediction model for each industry from training data using logistic-regression algorithm

No.	Variable name	Variable abbreviation	Beta Coefficient		
			Design	Fashion	Technology
	Constant value		1.430	1.716	1.586
1	Experience to launch project	EA1	2.008	1.935	1.553
2	Experience to back other project	EA2	2.910	4.157	3.118
3	No. of images	PA1	1.928	0.740	1.783
4	Having a project video	PA2	3.832	4.168	4.072
5	Having a project website	PA3	0.391	0.555	0.388
6	No. of available reward levels	PA4	2.084	3.540	2.808
7	Campaign duration	CA1	-0.801	-0.718	-0.903
8	Funding goal	CA2	-13.056	-13.763	-11.919
9	Delivery time of reward	CA3	-0.088	1.269	-0.637

method in order to divide the data into two parts: (a) training data in the proportion of 70 percent to be used as data for model training and construction and (b) testing data in the proportion of 30 percent to be used for model validation in the model development process. In the use of training data for model development, the researcher used a simple sample selection method (sample random sampling) in each industry to provide the number of successful and unsuccessful reward-based crowdfunding projects for each industry. There were the same number of examples for both successful and unsuccessful crowdfunding projects to develop a model of reward-based crowdfunding for each industry.

In developing models for each industry, researchers developed a model of a reward-based crowdfunding model using a machine-learning technique with supervised learning in a classification model that shows the project successfully raises funds, along with projects that have not been successful in crowdfunding by using the logistic-regression algorithm to develop a model that helps to predict the chance of success in a reward-based crowdfunding campaign for each industry. The models for each industry created using machine learning and the logistic-regression algorithm can be summed up as shown in Table 8.

Entrepreneurs' prior experience with crowdfunding and their prior backing of other crowdfunding projects are the independent variables in terms of the entrepreneurial factor for crowdfunding in the design, fashion, and technology industries. The coefficients of these variables in each industrial variable range from 1.553 to 2.008 and from 2.910 to 4.157, respectively. Similar to all three industries, the likelihood of success in a crowdfunding campaign positively correlates with the number of projects launched by entrepreneurs. This is because entrepreneurs can use the lessons learned from previous campaigns to strengthen current campaigns. This raises the possibility of crowdfunding success. Additionally, entrepreneurs should give funds to the crowdfunding efforts of other project creators to foster goodwill among backers and promote interest in the crowdfunding campaign from those backers.

The campaign image, campaign video, project website to display additional project specifics, and the number of various reward possibilities are the independent factors in the project factor of the design, fashion, and technology industries. In each business, the coefficients of each variable range from 0.740 to 1.928, 3.832 to 4.168, 0.3884 to 0.555 and 2.084 to 3.540, respectively. The success of reward-based crowdfunding projects positively correlates with all four independent factors across the three industries.

Therefore, when proposing a crowdfunding campaign, a campaign image should be used to help explain the project, as well as having a video and website of the project to show more project details and rewards and to give backers a full understanding and visibility of the project as well as

Table 9. Efficiency results of prediction model for each industry from training data using logistic-regression algorithm

Unit: No. of project	Industry		
	Design	Fashion	Technology
Total training projects	16,018	16,018	16,018
Expected result is true for			
Unsuccessful projects	6,628	6,985	6,880
Successful projects	7,288	7,197	7,514
Total	13,916	14,182	14,493
Accuracy rate (%)	*86.9%*	*88.5%*	*89.9%*

understanding of the nature of the rewards that will be well received. These also help to build the credibility of entrepreneurs and create backers' confidence in crowdfunding campaigns and good understanding of the nature of rewards to be received, which help to speed-up backers' funding decision. Entrepreneurs should also provide a range of reward options so that backers may more readily select incentives that are relevant and meet their demands.

In the design, fashion, and technology industries, independent variables for campaign duration and funding goal from campaign factor have coefficients in a narrow range between -0.718 to -0.903 and between -11.919 to -13.056, respectively. There is a negative correlation with the likelihood of success in reward-based crowdfunding campaigns. Therefore, entrepreneurs should set appropriate timing for crowdfunding. The number of days for fundraising should not be set too long because a crowdfunding campaign with a long campaign period does look interesting from the backers' point of view. Additionally, entrepreneurs should set a realistic funding goal based on the project's size and reasonable funding targets because there is a lower possibility of reaching the financial goal if the project's funding target is set too high.

But in the design and technology industries, the variables for reward delivery time have coefficients of -0.088 and -0.637. Entrepreneurs should therefore not set the time for delivery of rewards for too long. Such a variable for the fashion industry has a coefficient of 1.269 because the backers in the fashion industries can wait longer for delivery and use than those in the design and technology industries, which are more prone to obsolescence.

To evaluate how well the prediction model performs in each industry's reward-based crowdfunding, as shown in Table 9, this study employed the confusion matrix and discovered that the model's accuracy is approximately 86.9% for the design industry, 88.5% for the fashion business, and 89.9% for the technology industry.

Model's Validity Test

In the validity test, the researcher tested the validity of the model for evaluating the success of reward-based crowdfunding. The results of the validity testing of the model are summarized.

Internal Validity Test by Using Testing Data

Results of Internal Validity Test of Model From Testing Data Using Logistic-Regression Algorithm. The researcher separated the testing data into 30% of the 68,648 data samples to be utilized as data for testing the model's validation. The confusion matrix was applied by the researcher to examine the model's internal validity test findings. With the help of testing data, it is discovered that the model's accuracy from an internal validity test is 88.0%, as shown in Table 10.

Results of Internal Validity Test of Model From Testing Data Using Decision-Tree Algorithm. The internal validity test of the model was examined using the confusion matrix. With the testing

Table 10. Efficiency results of internal validity test from testing data using logistic-regression algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	8,725	42.4%	1,571	7.6%	10,296
	Successful campaign	906	4.4%	9,390	45.6%	10,296
	Total	9,631	46.8%	10,961	53.2%	20,592

Note. Remark: Accuracy rate is (8,725 + 9,390) / 20,592 = 88.0%.

Table 11. Efficiency results of internal validity test from testing data using decision-tree algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	8,835	42.9%	1,461	7.1%	10,296
	Successful campaign	1,045	5.1%	9,251	44.9	10,296
	Total	9,880	48.0%	10,712	52.0%	20,592

Note. Remark: Accuracy rate is (8,835 + 9,251) / 20,592 = 87.8%.

Table 12. Efficiency results of internal validity test for each industry from testing data using logistic-regression algorithm

Unit: No. of project	Industry		
	Design	Fashion	Technology
Total testing projects	6,864	6,864	6,864
Expected result is true for			
Unsuccessful projects	2,841	2,977	2,922
Successful projects	3,138	3,068	3,244
Total	5,979	6,045	6,166
Accuracy rate (%)	*87.1%*	*88.1%*	*89.8%*

data of the model, it is found that the internal validity test of the model has an accuracy value of 87.8%, as detailed in Table 11.

Results of Internal Validity Test of Model for Each Industry From Testing Data Using Logistic-Regression Algorithm. Confusion matrix was used to analyze the internal validity test of the model for the design industry. With the use of the model's testing data, it is discovered that the internal validity test's accuracy value is 87.1% for the design industry, 88.1% for the fashion industry, and 89.8% for the technology industry, as shown in Table 12.

External Validity Test by Using Separated Testing Data

Results of External Validity Test of Model From Separated Testing Data Using Logistic-Regression Algorithm. Three industries were included in the sample of 70,448 completed reward-based crowdfunding projects, which included both successful and unsuccessful campaigns. The researcher divided the data for testing of the model by extracting the number of samples from the total

Table 13. Number of reward-based crowdfunding projects for external validity testing

Industry	No. of reward-based crowdfunding projects		
	Successful projects	Successful projects	Total
Design	300	300	600
Fashion	300	300	600
Technology	300	300	600
Total	900	900	1,800

Note. Source: https://www.kickstarter.com.

Table 14. Efficiency results of external validity test from separated testing data using logistic-regression algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	769	42.7%	131	7.3%	900
	Successful campaign	85	4.7%	815	45.3%	900
	Total	854	47.4%	946	52.6%	1,800

Note. Remark: Accuracy rate is (769 + 815) / 1,800 = 88.0%.

sample data as separate data for another data set of 1,800 samples, which is a test using retrospective data for testing the external validity of the model developed from training data.

In dividing the data for the external validity test of the model using a multistage sampling method, the data are divided into each of the three industries, and then the desired sample determined in each industry equal to 600 samples per industry as external validation data, and then simple sample random sampling applied in each industry to obtain the number of successful and unsuccessful projects in crowdfunding campaigns to obtain the number of samples of each subgroup being equal, 300 samples per sub-group, as detailed in Table 13.

For the examination of the model external validity test findings, the researcher employed the confusion matrix. The model's external validity test is determined to have an accuracy value of 88.0% using data from 1,800 samples for external validation, as shown in Table 14.

Results of External Validity Test of Model From Separated Testing Data Using Decision-Tree Algorithm. The confusion matrix was utilized to analyze the outcomes of the model testing. The accuracy result from the model's external validation test is determined to be 87.1%, as shown in Table 15, using external testing data.

Results of External Validity Test of Model for Design Industry From Separated Testing Data Using Logistic-Regression Algorithm. The confusion matrix was utilized to analyze the outcomes of the model testing. With the help of external testing data, it is discovered that the accuracy value for the model's validation is 84.3% for the design industry, 89.5% for the fashion industry, and 89.7% for the technology industry, as shown in Table 16.

External Validity Test by Using New Testing Data

Results of Internal Validity Test of Model From New Testing Data Using Logistic-Regression Algorithm. To test the external validation of the model with new testing data that uses prospective data, the researcher collected additional secondary data of reward-based crowdfunding projects from the Kickstarter website. The crowdfunding campaign data were gathered from three industries: design

Table 15. Efficiency results of external validity test from separated testing data using decision-tree algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	767	42.6%	133	7.4%	900
	Successful campaign	100	5.6%	800	44.4%	900
	Total	867	48.2%	933	51.8%	1,800

Note. Remark: Accuracy rate is (767 + 800) / 1,800 = 87.1%.

Table 16. Efficiency results of external validity test for each industry from separated testing data using logistic-regression algorithm

Unit: No. of project	Industry		
	Design	Fashion	Technology
Total testing projects	600	600	600
Expected result is true for			
Unsuccessful projects	235	272	261
Successful projects	271	265	277
Total	506	537	538
Accuracy rate (%)	*84.3%*	*89.5%*	*89.7%*

industry, fashion industry, and technology industry, which were reward-based crowdfunding projects that had 3,763 completed funding campaign projects from February 2020 to October 2020. All project data were separated into the three industries: design (1,756 projects, or 47%), fashion (787 projects, or 21%), and technology (1,220 projects, or 32%), which contained information on both observable variables that would be used to develop the model and the crowdfunding results.

The confusion matrix was utilized by the researcher to analyze the outcomes of the model testing. It is discovered that the model validation result improves with the new data validation of 3,763 samples. The accuracy value is 90.6% as shown in Table 17.

Results of Internal Validity Test of Model From New Testing Data Using Decision-Tree Algorithm. The confusion matrix was utilized to analyze the outcomes of the model testing. With new

Table 17. Efficiency results of external validity test from new testing data using logistic-regression algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	1,062	28.3%	99	2.6%	1,161
	Successful campaign	254	6.7%	2,348	62.4%	2,602
	Total	1,316	35.0%	2,447	65.0%	3,763

Note. Remark: Accuracy rate is (1,062 + 2,348) / 3,763 = 90.6%.

Table 18. Efficiency results of external validity test from new testing data using decision-tree algorithm

Logistic regression algorithm		Expected result				
		Unsuccessful campaign		Successful campaign		Total
		No. of projects	% of total project	No. of projects	% of total project	No. of projects
Actual result	Unsuccessful campaign	1,071	28.5%	90	2.4%	1,161
	Successful campaign	286	7.6%	2,316	61.5%	2,602
	Total	1,357	36.1%	2,406	63.9%	3,763

Note. Remark: Accuracy rate is (1,071 + 2,316) / 3,763 = 90.0%.

Table 19. Efficiency results of external validity test for each industry from separated testing data using logistic-regression algorithm

Unit: No. of project	Industry		
	Design	Fashion	Technology
Total new testing projects	1,756	787	1,220
Expected result is true for			
Unsuccessful projects	297	160	607
Successful projects	1,287	521	541
Total	1,584	581	1,148
Accuracy rate (%)	*90.2%*	*86.5%*	*94.1%*

testing data validation of 3,763 samples, it is found that the model validation result has an accuracy value at 90.0% as shown in Table 18.

By analyzing the results of model testing from both logistic-regression and decision-tree methods, the model developed by the logistic-regression method can be used to explain to general users so as to understand easily, which helps to indicate the direction of the relationship of each factor as well. In comparison to the model generated using the decision-tree method, the model developed using the logistic regression method has a higher accuracy value from both internal and external validation testing.

Results of Internal Validity Test of Model for Design Industry From New Testing Data Using Logistic-Regression Algorithm. The researcher used the confusion matrix for the analysis of the model test results. It is discovered that the model validation result has an accuracy value of 90.2% for the design industry, 86.5% for the fashion industry, and 94.1% for the technology industry, as shown in Table 19.

DISCUSSION

Crowdfunding platforms produce a public more prepared for innovation and complex challenges, providing circular economy entrepreneurs a test environment to validate their ideas (Brown et al., 2017). Crowdfunding backers represent, in fact, a target of clients already averse to risk. Project initiators test the launch of innovative products within a market sample, also having the opportunity to establish a direct link with their early adopter (Stanko & Henard, 2017).

This study has added to the body of knowledge on reward-based crowdfunding by developing a thorough prediction model to evaluate the success of crowdfunding campaigns. The results of this study apparently demonstrate the crowdfunding prediction model and important independent variables that lead to success in a reward-based crowdfunding project. This research took the variables that are important to the success of reward-based crowdfunding from the literature and used them to develop a prediction model for success in reward-based crowdfunding campaigns. Secondary data of completed reward-based crowdfunding projects were collected from the Kickstarter website. The model was developed by a machine-learning technique with a supervised-learning approach in the classification model to predict the possibility of success in reward-based crowdfunding campaigns. The models were constructed using logistic-regression and decision-tree algorithms with collected data from three industries: design, fashion, and technology.

The model validation result using the logistical-regression algorithm is 88.2%, the accuracy value for the model internal validity test using the testing data is 88.0%, the accuracy value for the external validation test result using separated testing data is 88.0%, and the accuracy value for the test result for external validation using new testing data is 90.6%. For the model developed by the decision-tree algorithm, the accuracy of model validation is 88.8% and the results of the model validity test using testing data are 87.8%. External validation by separated testing data has a validity value of 87.1% and external validation by new testing has an accuracy value of 90.0%. As a result, the external validation test using new testing data of 3,763 crowdfunding campaigns gathered from a later data collection period from February 2020 to October 2020 suggests that this model for predicting crowdfunding outcomes might be applied to forecast future crowdfunding outcomes.

This research also employed a logistic-regression algorithm to construct a prediction model for reward-based crowdfunding success for each industry, which includes the design, fashion, and technology industries. The accuracy values of the model validation are equal to 86.9% in the design industry model, 88.5% in the fashion industry model, and 89.9% in the technology industry model, and the result of testing the internal validity of the model using testing data has an accuracy value equal to 87.1% 88.1. % and 89.8%, respectively, and the external validation test result with separated testing data has accuracy values at 84.3%, 89.5%, and 89.7%, respectively. External validation by new testing data has accuracy values at 90.2%, 86.5%, and 94.1%, respectively.

The prediction models from this study are validated by the literature, and the the results in this research outperform and are able to achieve greater levels of accuracy than the researchers discovered in the literature from Greenberg et al. (2013), Etter et al. (2013), and Hussaing et al. (2018), which provided an accuracy of 68%, 76%, and 80.4%, respectively.

The study's findings contribute to our knowledge and comprehension of reward-based crowdfunding campaigns. This is an innovative alternative to start-up entrepreneurs' fundraising models and provides important behavioral details and information for project creators to consider when deciding whether to run a crowdfunding campaign. Entrepreneurs can collect information about the public's valuation of their projects, reducing the entrepreneur's uncertainty before a new product's release on the market (Da Cruz, 2018). Project creators should consider crowdfunding initiatives as a test environment for testing consumer behavior when testing the quality and reliability of their product. This can be extended further into a body of knowledge to research how funders behave when supporting crowdfunding projects.

CONCLUSION

Due to the riskier and newly invented initiatives for entrepreneurs, reward-based crowdfunding is an alternative fundraising channel for those who need seed cash to assist the development of their prototypes. When a creator presents a project or product appeal at the launch of a reward-based crowdfunding campaign, funders lack information about the quality, requirements, and reputation of the project (Skirnevskiy et al., 2017). As a result, entrepreneurs still face difficulties when trying to raise money from a crowd. The many existing approaches studied factors influencing successful crowdfunding campaigns by mainly focusing on developing an estimation model using a traditional quantitative approach, such as descriptive analysis, multiple regression, logistic regression, robust linear regression, or employing a qualitative approach from interviews with crowdfunding experts on crowdfunding creators' and backers' perspectives toward crowdfunding campaigns. This study addresses this gap by developing a model for predicting the success of reward-based crowdfunding projects by machine-learning techniques including logistic-regression and decision-tree algorithms that have been employed. The data sets were retrospectively collected from historical records of campaigns on the Kickstarter website, one of the most well-known crowdfunding platforms on the internet and were divided into two parts, which consisted of training data for model training and construction and testing data for model validation in the model development process. To support this prediction model, extensive tests were carried out, and machine-learning methods, including logistic regression and decision tree algorithms, were applied.

The study's findings show that the logistic-regression and decision-tree models have accuracy rates of 88.2% and 88.8%, respectively. We reveal consequential considerations when entrepreneurs create reward-based crowdfunding strategies, especially when engaging with funders using online environments. The results thus offer valuable theoretical insights and practical guidance for reward-based crowdfunding platforms and entrepreneurs seeking to understand the funders' decision-making process that backs the crowdfunding campaigns.

Theoretical Contributions

The study represents a marked departure from previous research by extending the development of prediction models for three specific industries, which are design, fashion, and technology. More specifically, we offer insights for researching the link between the design, fashion, and technology industries and how they can reach funding targets in crowdfunding campaigns, especially reward-based models. The data sets from these three industries were separately constructed for each model in each industry to predict the probability of success in crowdfunding campaigns in design, fashion, and technology. To estimate the prediction models, extensive tests were carried out, and machine-learning methods including logistic regression have been employed to construct prediction models in each industry. Therefore, this study examines crowdfunding campaigns in different industries and explores the variations among them. These findings emphasize the importance of comparative analysis among the design, fashion, and technology industries, offering implications for the future. The results of prediction models in each industry can help entrepreneurs to reduce some risks connected to the creation of crowdfunding campaigns in the design, fashion, and technology industries, such as uncertainty in the market and the long-range perspective for creators and customers. In addition, it supports future entrepreneurs in evaluating the feasibility of the entrepreneurial journey.

Practical Implications

The results of this research provide several practical implications for crowdfunding platforms and entrepreneurs. First, this study provides practical insights for entrepreneurs, particularly in the context of the design, fashion, and technology industries to apply the proposed prediction model results as a guideline to prepare before the project is actually launched for a crowdfunding campaign and help to identify the most influential topical features embedded in project campaigns and hence to better promote their projects and improve the chance of raising sufficient funds for their projects. Second, according to the importance of creators' experience in launching crowdfunding projects, entrepreneurs could translate their online and offline experience and reputations into online crowdfunding contexts demonstrating competency for creating reward-based crowdfunding projects (Chalençon et al., 2017). Third, this research helps crowdfunding platforms, which are intermediaries that offer crowdfunding services online. By understanding the key factors that affect the success of reward-based crowdfunding, these platforms can design and improve the way that important information is presented on their websites so that it is more engaging and encourages backers to make a funding contribution. Fourth, to expand the availability of consulting services for the entrepreneurs who are their clients, crowdfunding platforms may also make use of the crowdfunding prediction model that the researcher has developed. Last, and even more interesting than understanding the potential of crowdfunding campaigns, policymakers should encourage an alternative form of finance explaining the nonmonetary value that entrepreneurs should achieve from this experience. At the same time, they should consult the crowdfunding environment to find new challenges to embrace, contributing to building a more adequate societal and environmental proposition.

Limitations and Future Research Directions

This study still has limitations and suggests the direction for future research. First, the data set used in this research focuses only on reward-based crowdfunding initiatives. Other types of crowdfunding projects, like equity-based and lending-based crowdfunding initiatives, could be a valuable area and should be studied in future research. In this vein, researchers interested in crowdfunding models should extend the research by analyzing other forms of crowdfunding environment such as equity-based or lending-based models.

Second, this study collected historical data on reward-based crowdfunding campaigns in three industries, which consisted of design, fashion, and technology industries. Therefore, the developed model in each industry can be used to evaluate the success of crowdfunding campaigns. However, the prediction model for other industries could be developed using more required data of crowdfunding projects in those industries. Academics can progress to exploring the value of crowdfunding for other specific industries in management studies.

Third, due to the limitations in the crowdfunding platform's disclosure policy for previous campaigns, the researcher may not have been able to gather data for all historical campaigns on the Kickstarter website, which is the crowdfunding platform used for this study. In light of this, if all available historical information on long-gone crowdfunding projects can be gathered, it might be possible to examine the variables affecting the success of reward-based crowdfunding campaigns and develop prediction crowdfunding models over different periods.

Fourth, this study considers features that are mainly numerical in nature. It can be extended to analyze textual data such as project description, updates, comments, and visual data, for example, images and videos. In the future, a model with multiple analytics, which also considers the textual and visual content of project, social media, and communication content, could be developed.

Last, the study does not address the potential variations across different cultures of backers' decision-making perspectives. Future research could conduct comparative cross-cultural analysis using data from different countries to explore these differences.

PROCESS DATES

07, 2024

This manuscript was initially received for consideration for the journal on 04/02/2024, revisions were received for the manuscript following the double-anonymized peer review on 07/07/2024, the manuscript was formally accepted on 06/26/2024, and the manuscript was finalized for publication on 07/08/2024

CORRESPONDING AUTHOR

Correspondence should be addressed to Somboon Prasobpiboon; somboon_mac@yahoo.com

REFERENCES

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review*, *37*(3), 355–375. 10.5465/amr.2010.0146

Ahsan, M., & Musteen, M. (2021). International opportunity development on crowdfunding platforms: A spatial, temporal, and structural framework. *International Business Review*, *30*(6), 101912. 10.1016/j.ibusrev.2021.101912

Banerjee, S., & Bose, I. (2022). An attractive proposition? Persuading retail consumers to prefer reward-based crowdfunding for owning upcoming technologies. *Information & Management*, *59*(6), 103663. 10.1016/j.im.2022.103663

Bannerman, S. (2013). Crowdfunding culture. *Journal of Mobile Culture*, *7*(1), 1–30.

Beaulieu, T., Sarker, S., & Sarker, S. (2015). A conceptual framework for understanding crowdfunding. *Communications of the Association for Information Systems*, *37*(1), 1–31.

Beier, M., & Früh, S. (2020). Reward-based crowdfunding as a tool to constitute and develop collaborative innovation networks (COINs). In *Digital transformation of collaboration:Proceedings of the 9th international COINs conference* (pp. 279-293). Springer International Publishing.

Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, *29*(5), 585–609. 10.1016/j.jbusvent.2013.07.003

Belleflamme, P., Omrani, N., & Peitz, M. (2015). The economics of crowdfunding platforms. *Information Economics and Policy*, *33*, 11–28. 10.1016/j.infoecopol.2015.08.003

Bennouna, G., & Tkiouat, M. (2019). Scoring in microfinance: Credit risk management tool–case of Morocco. *Procedia Computer Science*, *148*, 522–531. 10.1016/j.procs.2019.01.025

Berry, W. D., & Feldman, S. (1985). Multiple regression in practice. *Sage (Atlanta, Ga.)*, *50*, 43.

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. 10.1016/j.jbusres.2016.10.001

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

Brown, T. E., Boon, E., & Pitt, L. F. (2017). Seeking funding in order to sell: Crowdfunding as a marketing tool. *Business Horizons*, *60*(2), 189–195. 10.1016/j.bushor.2016.11.004

Cavalcanti Junqueira, M. I., & Soetanto, D. (2022). Funding decisions and the role of trust: A qualitative study of reward-based crowdfunding in the creative industries. *Management Decision*, *60*(8), 2174–2194. 10.1108/MD-08-2020-1095

Chalençon, L., Colovic, A., Lamotte, O., & Mayrhofer, U. (2017). Reputation, e-reputation, and value-creation of mergers and acquisitions. *International Studies of Management & Organization*, *47*(1), 4–22. 10.1080/00208825.2017.1241086

Chandna, V., & Salimath, M. S. (2018). Peer-to-peer selling in online platforms: A salient business model for virtual entrepreneurship. *Journal of Business Research*, *84*, 162–174. 10.1016/j.jbusres.2017.11.019

Cordova, A., Dolci, J., & Gianfrate, G. (2015). The determinants of crowdfunding success: Evidence from technology projects. *Procedia: Social and Behavioral Sciences*, *181*, 115–124. 10.1016/j.sbspro.2015.04.872

Cumming, D. J., Leboeuf, G., & Schwienbacher, A. (2020). Crowdfunding models: Keep-it-all vs. all-or-nothing. *Financial Management*, *49*(2), 331–360. 10.1111/fima.12262

Da Cruz, J. V. (2018). Beyond financing: Crowdfunding as an informational mechanism. *Journal of Business Venturing*, *33*(3), 371–393. 10.1016/j.jbusvent.2018.02.001

Daoyuan, S. (2016). *What is the recipe for success? An empirical analysis of crowdfunding project performance*. National University of Singapore.

Di Pietro, F., Prencipe, A., & Majchrzak, A. (2018). Crowd equity investors: An underutilized asset for open innovation in startups. *California Management Review*, *60*(2), 43–70. 10.1177/0008125617738260

Etter, V., Grossglauser, M., & Thiran, P. (2013, October). Launch hard or go home! Predicting the success of Kickstarter campaigns. In *Proceedings of the first ACM conference on online social networks* (pp. 177-182). 10.1145/2512938.2512957

Frydrych, D., Bock, A. J., Kinder, T., & Koeck, B. (2014). Exploring entrepreneurial legitimacy in reward-based crowdfunding. *Venture Capital*, *16*(3), 247–269. 10.1080/13691066.2014.916512

Gerber, E. M., & Hui, J. (2013). Crowdfunding: Motivations and deterrents for participation. [TOCHI]. *ACM Transactions on Computer-Human Interaction*, *20*(6), 1–32. 10.1145/2530540

Giones, F., & Brem, A. (2019). Crowdfunding as a tool for innovation marketing: Technology entrepreneurship commercialization strategies. In *Handbook of research on techno-entrepreneurship* (3rd ed., pp. 156–174). Edward Elgar Publishing.

Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013). Crowdfunding support tools: Predicting success and failure. In *CHI'13 extended abstracts on human factors in computing systems* (pp. 1815-1820). 10.1145/2468356.2468682

Guan, L. (2016). A short literature review on reward-based crowdfunding. In *2016 13th international conference on service systems and service management (ICSSSM)* (pp. 1-4). IEEE. 10.1109/ICSSSM.2016.7538636

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. John Wiley & Sons. 10.1002/0471722146

Hussain, N., Kamel, K., & Radhakrishna, A. (2018). Predicting the success of Kickstarter campaigns. *Cseweb. UCSD.edu*.

Jain, H., Khunteta, A., & Srivastava, S. (2020). Churn prediction in telecommunication using logistic regression and logit boost. *Procedia Computer Science*, *167*, 101–112. 10.1016/j.procs.2020.03.187

Kim, T., Por, M. H., & Yang, S. B. (2017). Winning the crowd in online fundraising platforms: The roles of founder and project features. *Electronic Commerce Research and Applications*, *25*, 86–94. 10.1016/j.elerap.2017.09.002

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). The Guilford Press.

Kunz, M. M., Bretschneider, U., Erler, M., & Leimeister, J. M. (2017). An empirical investigation of signaling in reward-based crowdfunding. *Electronic Commerce Research*, *17*(3), 425–461. 10.1007/s10660-016-9249-0

Kunz, M. M., Englisch, O., Beck, J., & Bretschneider, U. (2016). *Sometimes you win, sometimes you learn: Success factors in reward-based crowdfunding. Multikonferenz Wirtschaftsinformatik*. MKWI.

Kuswanto, H., Asfihani, A., Sarumaha, Y., & Ohwada, H. (2015). Logistic regression ensemble for predicting customer defection with very large sample size. *Procedia Computer Science*, *72*, 86–93. 10.1016/j.procs.2015.12.108

Laurell, C., Sandström, C., & Suseno, Y. (2019). Assessing the interplay between crowdfunding and sustainability in social media. *Technological Forecasting and Social Change*, *141*, 117–127. 10.1016/j.techfore.2018.07.015

Leone, D., Pietronudo, M. C., Gabteni, H., & Carli, M. R. (2023). Reward-based crowdfunding for building a valuable circular business model. *Journal of Business Research*, *157*, 113562. 10.1016/j.jbusres.2022.113562

Luo, X., Ge, L., & Wang, C. A. (2022). Crowdfunding for microfinance institutions: The new hope? *Management Information Systems Quarterly*, *46*(1), 373–400. 10.25300/MISQ/2022/15406

Maddala, G. S. (2001). *Introduction to econometrics* (3rd ed.). John Wiley & Sons.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, *29*(1), 1–16. 10.1016/j.jbusvent.2013.06.005

Müllerleile, T., & Joenssen, D. W. (2015). Key success-determinants of crowdfunded projects: An exploratory analysis. In *Data science, learning by latent structures, and knowledge discovery* (pp. 271–281). Springer Berlin Heidelberg. 10.1007/978-3-662-44983-7_24

Paschen, J. (2017). Choose wisely: Crowdfunding through the stages of the startup life cycle. *Business Horizons*, *60*(2), 179–188. 10.1016/j.bushor.2016.11.003

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in Large Margin Classifiers*, *10*(3), 61–74.

Sewaid, A., Garcia-Cestona, M., & Silaghi, F. (2021). Resolving information asymmetries in financing new product development: The case of reward-based crowdfunding. *Research Policy*, *50*(10), 104345. 10.1016/j.respol.2021.104345

Skirnevskiy, V., Bendig, D., & Brettel, M. (2017). The influence of internal social capital on serial creators' success in crowdfunding. *Entrepreneurship Theory and Practice*, *41*(2), 209–236. 10.1111/etap.12272

Song, C., Luo, J., Hölttä-Otto, K., Seering, W., & Otto, K. (2020). Crowdfunding for design innovation: Prediction model with critical factors. *IEEE Transactions on Engineering Management*, *69*(4), 1565–1576. 10.1109/TEM.2020.3001764

St. John, J., St John, K., & Han, B. (2022). Entrepreneurial crowdfunding backer motivations: A latent Dirichlet allocation approach. *European Journal of Innovation Management*, *25*(6), 223–241. 10.1108/EJIM-05-2021-0248

Stanko, M. A., & Henard, D. H. (2017). Toward a better understanding of crowdfunding, openness and the consequences for innovation. *Research Policy*, *46*(4), 784–798. 10.1016/j.respol.2017.02.003

Strzelecka, A., Kurdyś-Kujawska, A., & Zawadzka, D. (2020). Application of logistic regression models to assess household financial decisions regarding debt. *Procedia Computer Science*, *176*, 3418–3427. 10.1016/j.procs.2020.09.055

Szafraniec-Siluta, E., Zawadzka, D., & Strzelecka, A. (2022). Application of the logistic regression model to assess the likelihood of making tangible investments by agricultural enterprises. *Procedia Computer Science*, *207*, 3894–3903. 10.1016/j.procs.2022.09.451

Valanciene, L., & Jegeleviciute, S. (2013). Valuation of crowdfunding: Benefits and drawbacks. *Economics and Management*, *18*(1), 39–48. 10.5755/j01.em.18.1.3713

Wang, W., Liu, H., & Wu, Y. J. (2024). How do reward personalization options influence the public's willingness to participate in innovation projects? Insights from crowdfunding in Industry 5.0. *European Journal of Innovation Management*. Advance online publication. 10.1108/EJIM-07-2023-0584

Wang, W., Xu, Y., Wu, Y. J., & Goh, M. (2022). Linguistic information distortion on investment decision-making in the crowdfunding market. *Management Decision*, *60*(3), 648–672. 10.1108/MD-09-2020-1203

Yu, P. F., Huang, F. M., Yang, C., Liu, Y. H., Li, Z. Y., & Tsai, C. H. (2018, October). Prediction of crowdfunding project success with deep learning. In *2018 IEEE 15th international conference on e-business engineering (ICEBE)* (pp. 1-8). IEEE. 10.1109/ICEBE.2018.00012

Zahi, S., & Achchab, B. (2020). Modeling car loan prepayment using supervised machine learning. *Procedia Computer Science*, *170*, 1128–1133. 10.1016/j.procs.2020.03.055

Zhou, M., Lu, B., Fan, W., & Wang, G. A. (2018). Project description and crowdfunding success: An exploratory study. *Information Systems Frontiers*, *20*(2), 259–274. 10.1007/s10796-016-9723-129755287

Dr. Somboon Prasobpiboon received his Ph.D. in Technopreneurship and Innovation Management from Chulalongkorn University with a dissertation topic on the Innovative Reward-Based Crowdfunding Decision Model for Start-up Business. His research interests are broadly within the area of entrepreneurship, management and innovation including financial management

Assistant Professor Dr. Roongkiat is assistant professor at Chulalongkorn University. He received his Doctor of Philosophy in Pension Provision and Pension Fund Management from London School of Economics and Political Science, UK. His research focuses on pension, mutual fund, banking, microfinance.

Emeritus Achara Chandrachai, Ph.D. (Quantitative Business Analysis) Arizona State University. She is now Professor Emeritus of Chulalongkorn Business School, Chulalongkorn University, Thailand. Her research focuses on entrepreneurship, management and innovation.

Dr. Sipat Triukose is lecturer at Chulalongkorn University Big Data Analytics and IoT Center (CUBIC). He received his Doctor of Philosophy in Computer Science from Case Western Reserve University. His research focused on application of data science and digital technology in healthcare, large-scale performance evaluation and telemedicine.