Advances in Mathematical Finance
& Applications, 6(3), (2021), 408-423
DOI: 10.22034/AMFA.2020.1905837.1467

Published by IA University of Arak, Iran
Homepage: www.amfa.iau-arak.ac.ir

Modelling Crowdfunding Ensemble Learning Prediction

Mehran Saeidi Aghdam[a]*, *Akbar Alamtabriz*[b], *Alireza Bahiraie*[c], *Ahmad Sadeghi*[d]

[a]*Department of Eentrepreneurship, Qazvin Branch, Islamic Azad University, Qazvin, Iran*
[b]*Department of Industrial management, Shahid beheshti University, Tehran, Iran*
[c]*Department of Mathematics, Semnan University, Semnan, Iran*
[d]*Department of Geography, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran*

ARTICLE INFO
Article history:
Received 01 Auguest 2020
Accepted 08 October 2020

Keywords:
Prediction
Mathematical
Crowdfunding
Entrepreneurship

ABSTRACT
Crowdfunding is a new technology-enabled innovative process that is changing the capital market space. Internet-based applications, particularly those related to Web 2.0, have had a significant impact on sectors of society such as education, business, and medicine. The goal of this research is to fill a gap in the literature on mathematical modelling and prediction of ensemble learning in order to evaluate crowdfunding projects. The Mathematical model determines the cost of funding for the entrepreneur and the return investors will receive per period. A correct financial model is essential in order to keep all three stakeholders involved in the long term. The results show the designed model improved performance in predicting the evaluation of success or failure of Crowdfunding projects.

1 Introduction

Crowdfunding, a relatively new approach for raising capital for early-stage ventures, has grown by leaps and bounds in the last few years. Entrepreneurs launch a campaign on a web platform and solicit contributions from many potential backers. A primary way that entrepreneurs affect fundraising is by leveraging their social network, to drive traffic to their campaign [13]. Crowdfunding is believed to be derived from the broader concept of "crowdsourcing" and "micro-finance" [5, 22, 39]. The concept crowdsourcing was coined by Howe [24] in Wired Magazine referring to the practice of outsourcing organizations' tasks to a distributed network of individuals as an open call [8]. Micro-finance is a type of financial service that serves clients who have been excluded from the formal banking sector [18]. In line with these, crowdfunding utilizes the power of the crowd to finance small ventures, projects that cannot get access to institutional funders such as banks, credit institutions, or professional investors using crowdfunding platforms or social networks [24]. Crowdfunding means collecting funds from the crowd. Each investor provides a relatively small amount of financial support, acquiring the physical product or equity in return as Mollick [38] noted [23]. The astounding rise of crowdfunding over the last decade has raised awareness and interest in its potential [6]. Crowdfunding can be perceived as an approach [12], a process [19], and effort [18, 39], a capital formation strategy [22] or a financial mechanism [26] and all of these are valid as crowdfunding as it covers many uses and span across many fields [35]. Procurement of financial capital has long been accepted as one of the main difficulties

* Corresponding author. Tel.: *+989127845133*.
E-mail address: *Mehransaedi@gmail.com*

encountered by those starting new business ventures, and there exists a wealth of literature seeking to understand, analyze and make recommendations for businesses requiring capital [27, 28, 31, 32, 37]. Despite the attention, this area receives both academically and in practice, there remains a gap. New venture businesses, requiring early-stage or 'seed' funding are often left with two options: self-financing or external forms of capital. External finance in particular, including debt and equity, has attracted significant research attention in understanding their unique benefits and drawbacks [34]. Yet often there lacks the sufficient amount of either [33]. Forecast of financial time series is of the most critical issues in making financial decisions. Based on the economic events and data of the past, it provides a profitable method for the future. Financial time series forecasting is a challenging issue in the time-series field and has attracted many researcher's attention [20]. Forecasting in the financial time series is basically predicting the series behavior one or few steps ahead with the help of a number of variables. The variables used for forecasting are either economic or financial variables or in some cases, technical analysis output. In some studies, these variables are used in combination [17]. In recent years, different models have been used to predict demand by researchers, and since artificial intelligence techniques, including neural networks, genetic algorithms and fuzzy logic, have achieved successful results in solving complex problems (see [14] for more details). In 2014, the global crowdfunding industry market size was $16.2 billion. In 2015, it doubled to $34.4 billion, and it reached more than $50 billion in 2016. This financing model has maintained steady growth in recent years [44]. It is estimated that the size of the crowdfunding market will grow at an annual growth rate of approximately 30%, and the transaction volume is expected to reach $26 trillion by 2022. In terms of global distribution, the top five countries in terms of transaction volume are China, USA, UK, France, and Canada [41]. The problem of predicting success or failure not only involves the creator, but also the platforms; for this reason, it is increasingly common to direct the study towards the behavior of the platforms, as their benefits are proportional to the success of the projects. The problem of predicting success or failure not only involves the creator but also the platforms; for this reason, it is increasingly common to direct the study towards the behavior of the platforms, as their benefits are proportional to the success of the projects. One of the first important contributions using datasets in order to predict using artificial neural networks and supported by the theoretical framework known to date is a common behavior observed that in most successful projects, it has been observed that the projects that exceed the minimum amount of money for development. This research aims to develop a quantitative framework to understand and evaluate the quantitative and qualitative implications of the crowdfunding model for the entrepreneur and other stockholders. Through this end, a detailed understanding of crowdfunding is crucial. The focus will be on the Mathematical model that predicts the effectiveness of crowdfunding projects. Determining a correct Mathematical model is essential to restrain the negative consequences of asymmetrical information in order to keep all stakeholders involved in the long term.

2 Literature Review

2.1 Crowdfunding

Widespread Internet access, functioning social networking platforms together with the emancipation of the crowd propose interesting opportunities [29]. Leveraging these phenomena in a process called crowdfunding can help entrepreneurs gain the necessary start-up capital. Such a quest for alternative start-up capital is relevant as new ventures do not easily gain access to the necessary external finance at their early stages. In later periods, business angels and venture capital funds may fill gaps for larger amounts, however, costs for proofs-of-concept and the first entrepreneurial steps are often only financed by the entrepreneur, family, and friends [9]. Early debt-finance in such ventures is often brought up through a process identified in the literature as Bootstrapping [26]. Crowdfunding may thus provide a much-needed alternative for raising start-up capital for ventures seeking donations, debt, or equity finance. Crowdfunding is a constructed term that is often considered in the literature as project-based funding only and so the term in its current user does not fully comprise its full potential, which would also include more long-term commitments such as debt or equity shares [44]. Also, a distinct focus on donation-based crowdfunding for social entrepreneurs leaves out important market alternatives, where crowd members actually become shareholders. Especially equity-based Crowdfunding will thus inevitably cross the border of simple project financing [26].

Scholars see the roots of CF in a movement that has been labeled as crowdsourcing, which comprises using the crowd to obtain ideas, feedback, and solutions in order to develop corporate activities [8, 16, 19, 24]. A distinct feature of the ‘crowd’ is seen in literature as consisting of a large number of people, each contributing little, but with a possible high combined impact [7]. However, such a crowd is supposed to behave in unforeseen, chaotic, and complex manners [22]. Crowdfunding is a social entrepreneurship perspective can provide additional legitimacy to the venture, as the selection process by the crowd is perceived as democratic, and the crowd will thus select the social ideas it deems worthy and needed [7]. There are three main actors in crowdfunding projects, namely: Requesters, Funders, and Crowdfunding platforms. Researchers and crowdfunding platforms differ in their use of terminology, referring to people who request fund as “requesters” [25], “project founder” [36], “project initiators” [1], “creators” [15] or “creatives”, “designers”, “inventors”, etc. depending on platform’s focus [20], People who pledge funds may be referred to as “funders”, “fuelers”, “donors” “backers”, “investors” [7, 25]. In this section, for clarity, the author will refer to people who request funds as “requesters” and people who provide funds as “funders”. Crowdfunding platforms (websites) act as an intermediary, facilitating the transactions between the crowd of potential funders and the fund requesters [2]. Platforms usually require fees which can be platform fees, processor fees as a given percentage of fundraising [23].

2.2 Challenges of Crowdfunded Projects

To carry out ambitious projects or innovation through mass funding may be a not so obvious need for a creator, due to the skepticism caused by the direct loss of control, with the inclusion of new elements such as the uncertainty of non-compliance [21]. In addition to the intrinsic danger of using the internet as a medium, which involves accepting the challenge of navigating a very active environment susceptible to change, constantly stimulated by collective trends, updates, etc., this fact is also able to threaten projects with important innovation factor in their products, such as highly technological ones [41]. Part of the solution to these problems is mitigated by making use of the high capacity for sharing on blogs [32]. Blogs open the door to a multitude of users, from enthusiasts to specialists, who can provide solutions to various problems [27], and even generate valuable feedback during the development of the project, thus making it possible to improve the adaptation of the product to the environment and the user, complementing behavioral, technical and contextual skills [21].

On the other hand, sharing ideas or knowledge without any type of protection, intellectual or any other, increases the risk of plagiarism which is further aggravated if the project does not reach the minimum necessary amount, since it cannot be funded despite having been publicized both on the platform itself and on social networks, once the funding is canceled, it is not possible to present the project on the same platform, because one of the admission criteria is the originality of the projects [4].

2.3 Ensemble Learning

In a typical supervised learning setting, a set of instances, also referred to as a training set is given. The labels of the instances in the training set are known and the goal is to construct a model in order to label new instances. An algorithm that constructs the model is called inducer and an instance of an inducer for a specific training set is called a classifier. The main idea behind the ensemble methodology is to weigh several individual classifiers and combine them in order to obtain a classifier that outperforms every one of them. In fact, human being tends to seek several opinions before making any important decision. We weigh the individual opinions and combine them to reach our final decision [3]. De Condorcet was a French mathematician who among others wrote in 1785 the Essay on the Application of Analysis to the Probability of Majority Decisions. The ensemble idea in supervised learning has been investigated since the late seventies [47]. Tukey suggests combining two linear regression models. The first linear regression model is fitted to the original data and the second linear model to the residuals [11]. Two years later, Wolpert, suggested partitioning the input space using two or more classifiers [45]. The main progress in the field was achieved during the Nineties. Hansen and Salamon suggested an ensemble of similarly configured neural networks to improve the predictive performance of a single one [42].

At the same time Freund and Schapire laid the foundations for the award-winning Ada Boost algorithm by showing that a strong classifier in the probably approximately correct (PAC) sense can be generated by combining “weak” classifiers (that is, simple classifiers whose classification performance is only slightly better than random classification). Ensemble methods can also be used for improving the quality and robustness of unsupervised tasks. Ensemble methods can be also used for improving the quality and robustness of clustering algorithms [29]. Three ensemble-learning algorithms are employed in this study: bagging, random forest, and boosting. Bagging, developed by Breiman [10], is a machine-learning method that uses bootstrapping to create multiple training datasets from given datasets. The classification results generated using the data are arranged and combined to improve the prediction accuracy. Because the bootstrap samples are mutually independent, learning can be carried out in parallel [40]. Random forest, also proposed by Breiman [11], is similar to bagging. It is a machine-learning method in which the classification results generated from multiple training datasets are arranged and combined to improve the prediction accuracy [46]. However, whereas bagging uses all input variables to create each decision tree, random forest uses subsets that are random samplings of variables to create each decision tree. This means that the random forest is better suited than bagging for the analysis of high-dimensional data. Boosting is also a machine-learning method. Whereas bagging and random forest employ independent learning, boosting employs sequential learning [30]. In boosting, on the basis of supervised learning, weights are successively adjusted, and multiple learning results are sought. These results are then combined and integrated to improve overall accuracy. The most widely used boosting algorithm is Ada Boost, proposed by Freund and Schapire [43]. A neural network (NN) is a network structure comprising multiple connected units. It consists of an input layer, middle layer(s), and an output layer. The neural network configuration is determined by the manner in which the units are connected; different configurations enable a network to have different functions and characteristics. The feed-forward neural network is the most frequently used neural-network model and is configured by the hierarchical connection of multiple units. When the number of middle layers is greater than or equal to two, the network is called a deep neural network [29].

3 Methodology

To design Mathematical modeling and prediction of ensemble learning in order to evaluate crowdfunding projects on input data set some basic steps like Design, Data Collection and pre-processing, Model building and training, evaluation of prediction from those models are performed.

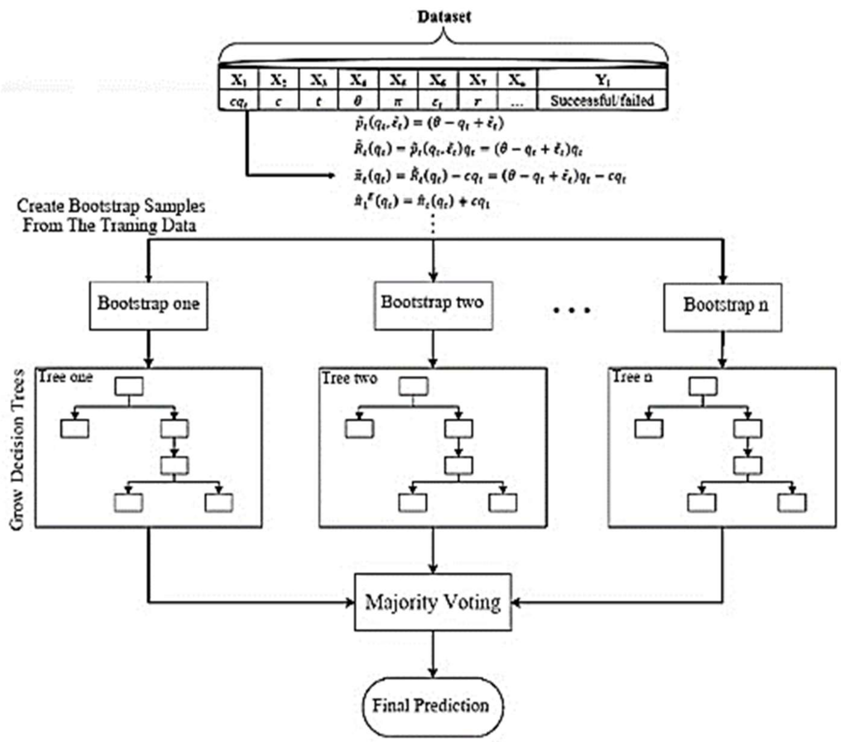

Fig. 1: The General Process of Research Model Experimental Procedure

This model based on Mathematical modeling crowdfunding undergoes these steps. Each of these steps is explained in this section. For experiment purposes, profit, number of production, income, market size, production costs, the price per unit, investor capital of daily trading data are predicted. The flow chart of the steps involved in the proposed method for each model is shown in Fig. 1.

4 Data and Models

In order to investigate the crowdfunding financial model, four different models are compared: The donation-based model, the pre-purchase model, the lending based model, and the equity crowdfunding model. The ventures are completely financed by one of these crowdfunding types and do not have other sources of funding. Therefore, the entrepreneur receives total production costs for the first period from investors. In this research have a data set includes features:

$$x_{i=}\ \{cq_t.c.t.\tilde{\varepsilon}_t.\tilde{R}_t(q_t).\tilde{\pi}_t(q_t).\tilde{p}_t(q_t.\tilde{\varepsilon}_t).r.\tilde{\pi}_t(q_t).\varepsilon_t.\beta.\pi.\theta.q_t\} \tag{1}$$

Where each item x_i has an associated class $y_i \in$ {1=successful, -1=failed}. Therefore, before weighting is based on dataset, assumptions are made and mathematical formulas are performed based on it in order to calculate profit, revenues, demand and the entrepreneur's value function. This model can be used to find the optimal granting financial credit to entrepreneurial projects in the crowdfunding system that maximizes expected profit for the beneficiaries and reduce investor risk. At the starting of the first period, t =1, the entrepreneur decides what quantity q_1. he will produce, according to his forecasted demand and the production costs per unit c. thereafter, he obtains an amount cq_1 from investors via crowdfunding platforms to produce q_1. In the second period, t = 2, the entrepreneur uses revenues made in the first period for the production of q_2. The price of the product is again assumed to be defined by a linear inverse demand function for the entrepreneur's product $\tilde{p}_t(q_t.\tilde{\varepsilon}_t) = (\theta - q_t + \tilde{\varepsilon}_t) \geq 0$. Here, θ represents the market size, and the random variable $\tilde{\varepsilon}_t$ is described by a normally distributed probability density function φ(.), with mean μ and variance δ^2, and cumulative distribution function Φ(.). By the end of each period, the entrepreneur realizes revenues from sales. Revenues after each period are described by:

$$\tilde{R}_t(q_t) = \tilde{p}_t(q_t.\tilde{\varepsilon}_t)q_t = (\theta - q_t + \tilde{\varepsilon}_t)q_t \tag{2}$$

These revenues result in a profit level for it:

$$\tilde{\pi}_t(q_t) = \tilde{R}_t(q_t) - cq_t = (\theta - q_t + \tilde{\varepsilon}_t)q_t - cq_t \tag{3}$$

Depending on the firm's profit in the first period, the entrepreneur pays out investors. Based on revenues, the entrepreneur determines the production quantity for the second period and uses internal cash for production. Hence, if internal cash is not sufficient to produce the optimal production quantity, the entrepreneur will produce less in the second period. Furthermore, if profit level $\tilde{\pi}_t(q_t)$ exceeds $\underline{\pi}$, the entrepreneur pays out a part β of the profit to investors. For now $\underline{\pi}$ is assumed to be zero, but his assumption is relaxed later on. Hence, if $\tilde{\pi}_t(q_t) = (\theta - q_t + \tilde{\varepsilon}_t)q_t - cq_t$ >0, the entrepreneur pays out to investors; otherwise he pays nothing to the investors in that particular period. After subtracting this payment, the firm remains with profit:

$$\tilde{\pi}_t(q_t) = \tilde{\pi}_t(q_t) - \beta\tilde{\pi}_t(q_t)I_{\{\tilde{\pi}_t(q_t)>0\}} \tag{4}$$

in period t. As β always lies between 0 and 1, the amount he owes to the investors is always smaller than the profit. Hence, the firm never defaults. Note that the entrepreneur received cq_1 from investors for production in the first period. In contrast to the debt case, the entrepreneur does not necessarily pay back the entire amount in that same period. Therefore, the entrepreneur's personal profits in the first period are given by:

$$\hat{\pi}_1{}^E(q_t) = \hat{\pi}_t(q_t) + cq_1 \tag{5}$$

Note that the assumption has been made that the entrepreneur aims to maximize his personal profits. Therefore, the value function for the entrepreneur can be expressed as follows:

$$\begin{aligned} & Max_{q_1 \geq 0} E_{\tilde{\varepsilon}_1}[\hat{\pi}_1(q_1) + cq_1 + V_2(q_1.\tilde{\varepsilon}_1)] \\ & \quad \text{Where } V_2(q_1.\tilde{\varepsilon}_1) = Max_{q_2 \geq 0} E_{\tilde{\varepsilon}_2}[\hat{\pi}_1(q_1)] \\ & \qquad\qquad \text{s.t.} \quad cq_2 \leq \hat{\pi}_1(q_1) + cq_1 \end{aligned} \tag{6}$$

In the next step initialize the weight for each data point as:

$$w(x_i.y_i) = \frac{1}{n} \tag{7}$$

For iteration $\mathrm{m} = 1, \dots, \mathrm{M}$.
Fit weak classifiers to the data set and select the one with the lowest weighted classification error**:**

$$J_m = \sum_{i=1}^{n} w_m^{(i)} \times I(y^{(i)} \neq h_m(x^{(i)})) \tag{8}$$

Calculate the weighted error of weak classifier:

$$\epsilon_m = \frac{\sum_{i=1}^{n} w_m^{(i)} \times I(y^{(i)} \neq h_m(x^{(i)}))}{\sum_{i=1}^{n} w_m^{(i)}} \tag{9}$$

And the new component is assigned votes base on its error:

$$\alpha_m = \frac{1}{n} \ln(\frac{(1 - \epsilon_m)}{\epsilon_m}) \tag{10}$$

For any classifier with an accuracy higher than 50%, the weight is positive. The more accurate the classifier, the larger the weight. While for the classifier with less than 50% accuracy, the weight is negative. It means that we combine its prediction by flipping the sign. The normalized weights are updated:

$$w_{m+1}^{(i)} = w_m^{(i)} e^{\alpha_m I(y^{(i)} \neq h_m(x^{(i)}))} \tag{11}$$

If a misclassified case is from a positive weighted classifier, the "$\boldsymbol{e}$" term in the numerator would be always larger than 1. Thus misclassified cases would be updated with larger weights after an iteration. The same logic applies to the negatively weighted classifiers. The only difference is that the original correct classifications would become misclassifications after flipping the sign. After M iteration can get the final prediction by summing up the weighted prediction of each classifier. Combined classifier:

$$\hat{y} = sign(h_M(x)) \text{where } h_M(x) = \sum_{m=1}^{M} a_m \, h_m(x) \tag{12}$$

This model refers to the procedures employed to train multiple learning machines and combine their outputs, treating them as a committee of decision-makers. The principle is that the committee decision, with individual predictions combined appropriately, should have better overall accuracy, on average, than any individual committee member. This model very often attains higher accuracy than single models. The members of this model might be predicting real-valued numbers, class labels, posterior probabilities, rankings, clustering's, or any other quantity. Therefore, their decisions can be combined by many methods, including averaging, voting, and probabilistic methods. The majority of ensemble learning methods are generic, applicable across broad classes of model types and learning tasks.

5 Numerical Experiments

Numerical experiments have been executed to assess the consequences of the Mathematical model. Subsequently, those experiments have been used to compare the models with each other. In this numerical analysis, the optimal operating policy for the entrepreneur who has obtained funding by crowdfunding is investigated. The experiments aim to obtain the optimal production quantity that maximizes the payoff for the entrepreneur over both periods and the resulting optimal profit for the entrepreneur. The optimal operating policy is compared to the monopoly quantity $\boldsymbol{q_m}$, which is the profit-maximizing production quantity when the projects would not be exposed to risk. This production quantity maximizes project value. In research, methodology consists of five steps.

Imports/Initial Data: In this research, for Initial data, used historical data for crowdfunding project from Kickstarter projects. It has 378658 projects from 2016/01/01 to 2020/04/20.

Table 1: Sample of Initial Data

Name	Market _size	Production	Production _qua	Launched	Pledged	Backers	Usd – goal	State
songs	1000	65	660.62	8/11/2016	0	0	1533.95	failed
Greeting	30000	137	68000	9/2/2017	2421	15	30000	Successful
Deployment	45000	587	7500	2/1/2016	220	3	45000	failed
Capital	5000	168	29700	3/17/2016	1	1	5000	failed
Community	19500	200	50000	7/4/2017	1283	14	19500	Canceled
Monarch	50000	781	3063.58	2/26/2016	52375	224	50000	successful

Mathematical Modelling: The experiments in the profit case aim to identify the optimal production quantity in the first period under the assumption that internal cash should be used for production in the second period. In addition, the expected two-period profit is obtained by the entrepreneur. The optimal production quantity in the second period has been defined as the monopoly quantity q_m. The entrepreneur needs to make sufficient profit in the first period to be able to produce this optimal quantity in the second period. The amount he needs to pay out to investors reduces profit, which can be used for production in the second period. Therefore, the influence of the amount that should be paid to investors is investigated. Namely, this amount affects the risk that the optimal production quantity cannot be produced in the second period. Experiments have been executed both for β= {0,0.1, 0.2, 0.3} and π= {0,50, 100}. In addition, the effects of the market parameter are considered, wherefore market size is varied; θ= {10,20,30}. Their individual impact and the managerial implications of their interactions are explored. The objective function for the profit has been defined, which should be maximized by the entrepreneur to gain maximum expected profit $Max_{q_1\geq 0}E_{\tilde{\varepsilon}_1}[\hat{\pi}_1(q_1)+cq_1+V_2(q_1.\tilde{\varepsilon}_1)]$.

First, the objective function is plotted against the production quantity in the first period q_1 for different payout percentages ß. An example for θ = 15 is illustrated in Fig. 2.

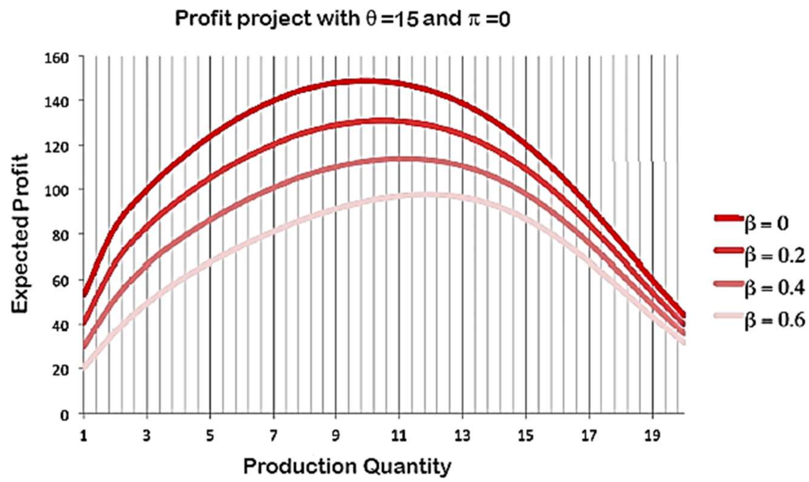

Fig. 2: Profit Projects with Different θ

The θ indicates the percentage of profit that the entrepreneur needs to pay to investors. If the entrepreneur does not have to pay to investors, there is no increased risk that the optimal production quantity cannot be produced in the second period. Hence, case θ = 0 can be seen as the benchmark. In that case, the entrepreneur produces around $\frac{\theta}{2}$ in the first period, which is significantly larger than the monopoly quantity $\frac{\theta-c}{2}$. This means that the entrepreneur overproduces in the first period (i.e. produces more than the monopoly quantity). The consequences of varying market size for optimal production quantity and optimal profit are demonstrated in Table 2.

Table 2: Optimal Production Quantity and Optimal Profit

	$\theta = 5$		$\theta = 10$		$\theta = 15$	
	$q_1{}^*(q_m)$	$\hat{\pi}$	$q_1{}^*(q_m)$	$\hat{\pi}$	$q_1{}^*(q_m)$	$\hat{\pi}$
ß=0	5.06	28.91	9.93	148.82	15.02	369.01
ß=0.2	5.21	27.32	9.96	131.01	15.73	314.62
ß=0.4	5.37	25.78	10.48	113.92	16.92	258.27
ß=0.6	5.53	24.26	11.82	97.64	18.84	208.32
ß=0.8	5.70	22.74	12.59	81.45	20.68	164.15

The experiments in the revenue case aim to investigate the consequences of the assumption that internal cash should be used for production in the period under revenue sharing conditions. In addition, the expected two-period profit is obtained for the entrepreneur. However, in contrast to the profit case, the optimal production quantity in the period is not the monopoly quantity. The entrepreneur needs to make sufficient profit in the first period to be able to produce this optimal quantity in the second period. The payout to investors reduces profit, which can be used for production in the second period. Hence, the influence of the amount that should be paid to investors is investigated, because it affects the risk that the optimal production quantity cannot be produced in the second period. Experiments have been executed both for payout percentage β={0,0.2, 0.4, 0.6} and threshold R= {0,50, 100}. In addition, the effect of the market parameter is considered, wherefore market size is varied; θ= {10, 20, 30). Their individual impact and the managerial implications of their interactions are explored. An example with θ = 20 and R = 0 is illustrated in Fig. 3.

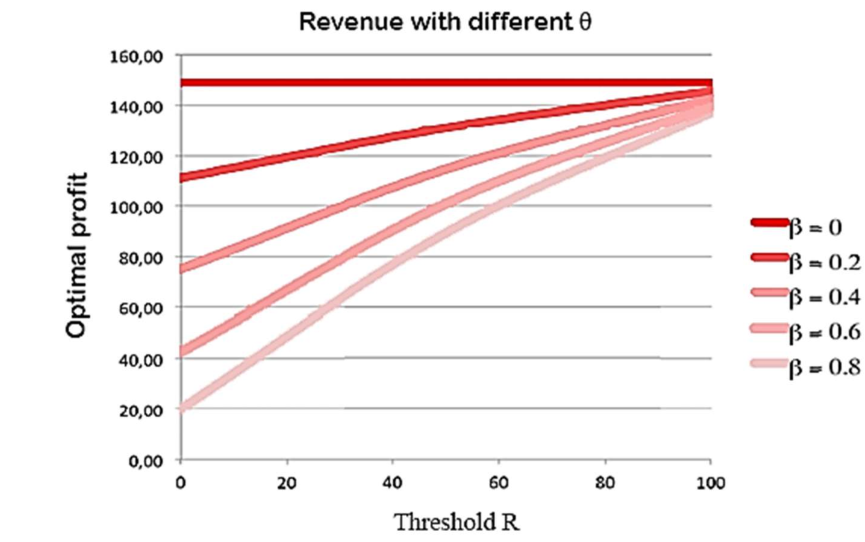

Fig. 3: Revenue Projects with Different θ

Table 3 shows the optimal production quantity in the first period with different values. Between brackets, the monopoly quantity is given.

Table 3: Revenue Optimal Production Quantity and Optimal Profit with Varying R and θ

	$R = 5$		$R = 10$		$R = 15$	
	$q_1{}^*(q_m)$	$\hat{\pi}$	$q_1{}^*(q_m)$	$\hat{\pi}$	$q_1{}^*(q_m)$	$\hat{\pi}$
ß=0	9.42	96.24	9.42	96.24	9.42	96.24
ß=0.2	9.58	87.35	9.47	95.23	9.72	142.64
ß=0.4	9.65	74.36	9.56	112.42	9.54	134.74
ß=0.6	9.78	52.73	9.68	90.33	9.33	138.24
ß=0.8	9.74	26.71	9.72	87.36	9.16	134.57

First the consequences of varying β are explored. Fig. 3 illustrates that the entrepreneur produces approximately in the first period for any value of β. In general, if R=0 the entrepreneur produces around $\frac{\theta}{2}$ in the first period and the production quantity increases slightly for increasing β. Result showed the optimal production quantity in the first period $q_1 *$ with different values of θ. Between brackets, the monopoly quantity is given. Furthermore, the implications of increasing the threshold R are investigated. Note that if revenue level $\tilde{\pi}_t(q_t) + cq_t$ exceeds R, the entrepreneur pays out a fraction β of $\tilde{\pi}_t(q_t) + cq_t - R$ to the investors. As R increases, the entrepreneur pays less to investors. If R becomes positive and fraction β is increased, this does affect the first-period production quantity with θ = 20 notably. The entrepreneur decreases production for increasing β in that case. With θ = 10 and θ = 30, increasing R has a reduced effect on the optimal production quantity in the first period. Moreover, increasing R does influence q_n. However, after a particular point, the production quantity in the second period does not decrease for higher β, because the threshold is not reached anymore.

Data Normalization: Normalization is changing the values of numeric columns in the dataset to a common scale, which helps the performance of model. In this research, normalization has been done by finding the numerical maximum in each criterion of the data set and dividing the input of each criterion to that maximum value.

Creating the Model: In this research implement the experiments using python, specifically, the “Sklearn” package for bagging, “Random Forest” for random forest, “Sklearn” package for boosting (AdaBoost algorithm), and “Kears” package for DNN. Furthermore, it analyzes the prediction accuracy rate of each method for two cases i.e., original and normalized data. Then, it examines the classification ability of each method based on the ROC curve, AUC value, and F-score.

Tables 4 and 5 report the results obtained using the original data. The tables show that boosting has the best performance and yields higher than 80% prediction accuracy rate on average, with a small standard deviation for both training and test data. None of the neural network models exceed an 80% average accuracy rate for test data. Thus, it is clear that boosting achieves a higher accuracy prediction than neural networks. The prediction accuracy rate for test data is less than 60% for bagging and random forest. In addition, the difference in ratios between training and test data does not have an obvious influence on the results of the analysis.

Table 4: Prediction Accuracy of Each Method for Original Data.

Method	Accuracy ratio of training data		Accuracy ratio of test data	
	Average	Standard Deviation	Average	Standard Deviation
Random Forest	82.16	0.004	72.25	0.008
Boosting	81.25	0.003	75.34	0.002
Bagging	80.96	0.015	73.67	0.814

Table 5: Performance of Models

Method			Accuracy ratio of training data		Accuracy ratio of test data	
Model	Activation function	Middle layer	Average	Standard Deviation	Average	Standard Deviation
DNN	Tanh	2	83.29	0.785	81.57	0.924
NN	Dropout	2	80.64	0.036	81.05	0.016

Fig. 4 displays ROC curves with AUC and F-score for the case using normalized data and the ratio between the training and test data of 70% to 30%. In the figure, sensitivity (vertical axis) corresponds to the true positive ratio, whereas 1—specificity (horizontal axis) corresponds to the false-positive ratio. The graphs indicate that the ROC curve for boosting and neural network models have desirable properties except for the case for the Tanh activation function with dropout. The AUC values and F-score are also shown for the figure. It is found that the highest AUC value is obtained for boosting (0.735). The highest F-score is also obtained for boosting (0.721). Thus, the classification ability to boost is superior to other machine-learning methods. This may be because boosting employs sequential learning of weights.

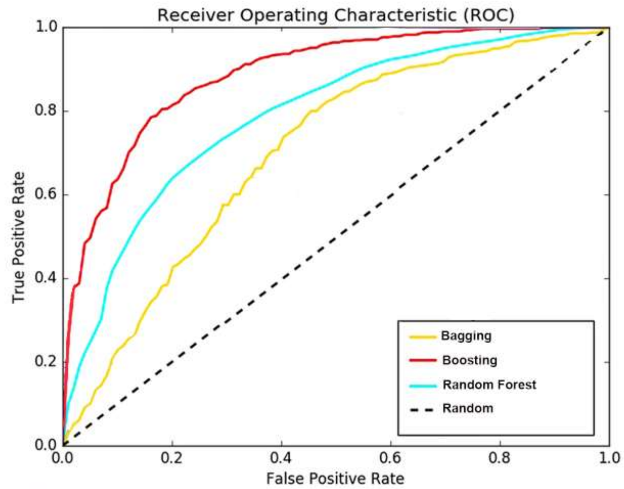

Fig. 4: Receiver Operating Characteristic (ROC) Curve for Bagging, Boosting and Random Forest

In this study, analyzed default data crowdfunding projects in Kickstarter and compared the prediction accuracy and classification ability of three ensemble-learning methods: bagging, random forest, and boosting, with those of various neural-network methods using two different activation functions. Results can be summarized as follows:

- The classification ability to boost is superior to other machine-learning methods.
- The prediction accuracy rate, AUC value, and F-score of NN are better than those of DNN when Tanh is used as an activation function.

6 Conclusion and Discussion

Crowdfunding is an emerging international financial activity often performed via an internet-mediated platform. With the rapid growth of this financial system, rising risks would influence a participant's decision making. In crowdfunding, backers and funders make decisions as a repeated game. Individual decision-makers interact and exchange information with each other through the crowdfunding process and other interaction assisted tools. He or she learns from others' experiences again and again to form his or her own decisions. Backers make decisions with the help of experience learned from feedback. Funders can also notice the reaction of backers when feedback is received by the backers. As an essential part of internet finance, crowdfunding plays an increasingly important role in today's economy. The success of crowdfunding depends on different factors. In recent years, more and more financing problems, such as Profit prediction, pricing, market prediction risk management, and revenue prediction, have been presented in the literature adopting formulation and solution approach rooted in mathematical theory, as well as optimization methods. Due to the acceptance of this financial method by stakeholders and the increasing speed of financing projects, credit risk assessment has attracted many research interests from both academic and industrial communities. A more accurate, consistent, and robust evaluation technique crowdfunding projects can significantly reduce future costs and prediction of successful or failed crowdfunding projects. In this study, a multistage neural network ensemble the learning model is proposed for successful or failed crowdfunding projects assessment. Different from commonly used ‘‘one-member-one-vote’’ or ‘‘majority-rule’’ ensemble, the novel neural network ensemble aggregates the decision values from the different neural ensemble members, instead of their classification results directly. The new ensemble strategy consists of two critical steps: scaling, which transforms decision values to degrees of reliability, and fusion, which aggregates degrees of reliability to generate final classification results.

Reference

[1] Allison, T.H., Davis, B.C., Webb, J.W., Short, J.C., *Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance*. J. Bus. Ventur. 2017, **32**, P. 707–725, Doi:10.1016/j.jbusvent.2017.09.002

[2] Anglin, A.H., Wolfe, M.T., Short, J.C., McKenny, A.F., Pidduck, R.J., *Narcissistic rhetoric and crowdfunding performance*, A social role theory perspective, J. Bus. Ventur. 2018, **33**, P. 780–812.
Doi: 10.1016/j.jbusvent.2018.04.004.

[3] Bauer, E., Kohavi, R., *An empirical comparison of voting classification algorithms: Bagging, Boosting, and variants,* Machine Learning, 1999, **36**, P. 105–139, Doi: 10.1023/A:1007515423169.

[4] Beaulieu, T., Sarker, S., *Discursive meaning creation in crowdfunding: A socio-material perspective,* Paper presented at the International Conference for Information Systems, Milan, Italy, 2013.
Doi: 10.1098/102515423169.

[5] Belleflamme, P., Lambert, T., Schwienbacher, A.,*Crowdfunding: Tapping the Right Crowd,* Journal of Business Venturing, 2014, **29**(5), P.585–609. Doi: 10.1016/j.jbusvent.2013.07.003.

[6] Belleflamme, P., Omrani, N., Peitz, M., *The economics of crowdfunding platforms,* Information Economics and Policy, 2015, **33**, P. 11–28, Doi: 10.1016/j.infoecopol.2015.08.003.

[7] Belleflamme, P., Lambert, T., Schwienbacher, A., *Digital Business Models: Understanding Strategies,* 2010, P. 1–30, Doi: 10.2478/fman-2018-0011.

[8] Boons, M., Stam, D., and Barkema, H.G., *Feelings of Pride and Respect as Drivers of Ongoing Member Activity on Crowdsourcing Platforms*, Journal of Management Studies, 2015, **52**(6), P.717–74.
Doi: 10.1111/joms.12140.

[9] Brabham, D.C., Ribisl. K.M., Kirchner. T. R., Bernhardt, J.M., *Crowdsourcing Applications for Public Health*, American Journal of Preventive Medicine, 2014, **46**(2), P. 179-187, Doi: 10.1016/j.amepre.2013.10.016

[10] Breiman, L., *Random forests. Machine Learning*, 2001, **45**(1), P. 5–32, Doi: 10.1023/A:1010933404324.

[11] Breiman, L., Bagging, Predictors Machine Learning,1996, **24**(2), P. 123–140, Doi: 10.1023/A:101805431435 0.

[12] Bruton, G.D., Khavul, S., Chavez, H., *Microlending in Emerging Economies: Building a New Line of Inquiry from the Ground Up,* Journal of International Business Studies, 2011, **42**(5), P. 718–739, Doi: 10.1057/jibs.2010.58.

[13] Burtch, G., Ghose, A., Wattal, S.*, The Hidden Cost of Accommodating Crowdfunder Privacy Preferences:A*

Randomized Field Experiment, Management Science, 2015, **61**(5), P. 949–962, Doi: 10.1287/mnsc.2014. 2069.

[14] Davoodi, A., Kasbi, I., Dadashi, K., *Stock price analysis using machine learning method (Non-sensory-parametric backup regression algorithm in linear and nonlinear mode),* Advances in Mathematical Finance and Applications, 2019, **4**(3), P. 1-13. Doi: 10.22034/AMFA.2019.1869838.1232.

[15] Dietterich, T.G., *Ensemble methods in machine learning, in Proc. Int. Workshop Multiple Classifier Syst*. Berlin, Germany: Springer, 2020, P. 1–15, Doi: 10.1007/3-540-45014-9_1.

[16] Dibachi, H., Behzadi, M.H., Izadikhah, M., *Stochastic Modified MAJ Model for Measuring the Efficiency and Ranking of DMUs*, Indian Journal of Science and Technology, 2015, **8**(8), P.1-7. Doi: 10.17485/ijst/2015/v8iS8/71505

[17] Farshadfar, Z., Marcel Prokopczuk, M., *Improving Stock Return Forecasting by Deep Learning Algorithm,* Advances in Mathematical Finance and Applications, 2019, **4**(3), P. 1-13. Doi: 10.22034/AMFA.2019.584494.1 173.

[18] Feller, J., Gleasure, R., Treacy, S., *From the wisdom to the wealth of crowds: A meta triangulation of crowd-funding research*, 2013, **1**(2), Doi: 10.1007/978-3-319-13817-6_36.

[19] Gerber, E. M., Hui, J. S., Kuo, P. Y., *Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms*, 2012, Doi: 10.1016/2012.05.033.

[20] Ghasemzadeh, M., Mohammad-Karimi, N., Ansari-Samani, H., Machine learning algorithms for time series in financial markets, Advances in Mathematical Finance and Applications, 2020, **5**(3), P.479-490. Doi: 10.22034/AMFA.2020.674946.

[21] Giudici, G., Guerini, M., Lamastra, C.R., *Why Crowdfunding Projects Can Succeed: The Role of Proponents Individual and Territorial Social Capital*, 2013, Doi: 10.2139/ssrn.2255944.

[22] Guo, K., Tang, Y., Zhang, P., *CSF: Crowdsourcing semantic fusion for heterogeneous media big data in the internet of things,* Trends in Food Science & Technology, 2017, **66**, P.63-72, Doi: 10.1016/j.inffus.2017.01.008.

[23] Hemer, J., Schneider, U., Dornbusch, F., Frey, S., *Crowd funding undand ere Formen670 informeller Mikrofinanzierung in der Projekt-und Innovationsfinanzierung,* 2011, ISBN-13: 978-3839603130.

[24] Howe, J., *Crowdsourcing: How the power of the crowd is driving the future of business, New York: Random House*, 2008. ISBN-13: 978-0307396211.

[25] Howe, J., *Crowdsourcing: A definition Crowdsourcing*. 2006, Doi: 10.1007/978-3-319-18341-1_3.

[26] Jenik, I., Lyman, T., Nava A., *Crowdfunding and Financial Inclusion. CGAP, US: World Bank Group.* 2017.

[27] Jones, O., Gatrell, C., *Editorial: The Future of Writing and Reviewing for IJMR." International Journal of Management Reviews*, 2014, **16**(3), P.249–264, Doi: 10.1111/ijmr.12038.

[28] Kleemann, F., Vob, G.G., Rieder, K., *Un(der)paid Innovators: The Commercial Utilization of Consumer Work Through Crowdsourcing*, Translated by S.S. Gissendanner. Science, Technology & Innovation Studies, 2008, **4**(1), Doi: 10.17877/DE290R-12790.

[29] Krogh, A., Vedelsby, J., *Neural network ensembles, cross validation, and active learning*. In Tesauro, G., Touretzky, D.S., Leen, T.K., eds.: Advances in Neural Information Processing Systems 7. MIT Press, Cambridge, 1995, P. 231–238.

[30] Kuncheva, L.I., Whitaker, C.J., *Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy*. Machine Learning, 2003, **51**(2), P. 181–207. Doi: 10.1023/A:1022859003006.

[31] Lambert, T., Schwienbacher, A., *An Empirical Analysis of Crowdfunding*, 2010.

[32] Larralde, B., Schwienbacher, A., *Crowdfunding of Small Entrepreneurial Ventures,* The Oxford Handbook of Entrepreneurial Finance, edited by D. Cumming, 2012, **369**. Doi: 10.2139/ssrn.1699183

[33] Lehner, O.M., *Crowdfunding Social Ventures: A Model and Research Agenda,* Venture Capital, 2013, **15**(4), P. 289–311, Doi: 10.1080/13691066.2013.782624.

[34] Ley, A., Weaven, S., *Exploring Agency Dynamics of Crowdfunding in Start-Up Capital Financing,* Academy of Entrepreneurship Journal, 2011, **17**(1), P. 85–110.

[35] Li, Y., Zhang, Z., Wang, R., Chen, Y., *Consumer Purchase Intention Toward Crowdfunding Products/Services: A Cost–Benefit Perspective*, Sustainability, 2019, **11**, 3579. Doi: 10.3390/su11133579.

[36] Lu, H., Zhou, H., Wang, Junxian, Wang, T., Dong, X., Zhuang, Z., Li, C., *Ensemble learning for independent component analysis of normal galaxy spectra*, The Astronomical Journal, 2006, **131**, P. 790-805. Doi: 10.1086/498711.

[37] Martínez-Climent, C., Costa-Climent, R., Oghazi, P, *Sustainable Financing through Crowdfunding. Sustainability*, 2019, **11**, P. 934-956. Doi: 10.3390/su11030934.

[38] Mollick, E., *The Dynamics of Crowdfunding: An Exploratory Study*, Journal of Business Venturing, 2014, **29**(1), P. 1–16, Doi: 10.1016/j.jbusvent.2013.06.005.

[39] Mollick, E., Robb, A., *Democratizing Innovation and Capital Access: The Role of Crowdfunding,* California Management Review, 2016, **58**(2), P. 72–87. Doi: 10.1525/cmr.2016.58.2.72

[40] Shravya, Ch., Pravalika, K., Subhani, Sh., *Prediction of breast cancer using supervised machine learning techniques*, Int. J. Innov. Technol. Explor. Eng, 2019. Doi: 10.1109/R10-HTC.2017.8288944

[41] Simons, A., Kaiser, LF., vom Brocke, J., *Enterprise crowdfunding: foundations*, *applications, and research findings*, Bus Inf SystEng, 2019, **61**, P. 113–121. Doi: 10.1007/s12599-018-0568-7.

[42] Testa, S., Nielsen, KR., Bogers, M., Cincotti, S., *The role of crowdfunding in moving towards a sustainable society,* Technol Forecast Soc Chang, 2019, **141**, P. 66–73. Doi: 10.1016/j.techfore.2018.12.011.

[43] Wang, S.-H., Zhang, Y.-D., Yang, M., Liu, B., Ramirez, J., Gorriz, J.M. *Unilateral sensorineural hearing loss identification based on double-density dual-tree complex wavelet transform and multinomial logistic regression.* Integr. Comput-Aided Eng, 2019, **26**(4), P. 411–426. Doi: 10.3233/ICA-190605.

[44] Yang, X., Zhao, K., Tao, X., Shiu, E, *Developing and Validating a Theory-Based Model of Crowdfunding*

Investment Intention—Perspectives from Social Exchange Theory and Customer Value Perspective, Sustainability, 2019, **11**, P. 25-56. Doi: 10.3390/su11092525.

[45] Zhou, L., Zhang, Z., Chen, Y.-C., Zhao, Z.-Y., Yin, X.-D., Jiang, H.-B, *A deep learning-based radiomics model for differentiating benign and malignant renal tumors*, Transl. Oncol, 2019, **12**(2), P. 292–300. Doi: 10.1016/j.tranon.2018.10.012.

[46] Zhou, Z.H., Jiang, Y., Chen, S.F, *Extracting symbolic rules from trained neural network ensembles*, AI Communications, 2003, **16**(1), P. 3–15. Doi: 10.1007/s11227-017-2228.

[47] Zhu, X., Zhang, C., Liang, G., *An empirical study of bagging predictors for different learning algorithms,* 25th Conference on Artificial Intelligence, 2011, P. 1802–1803. Doi: 10.1007/978-3-642-25853-4_26.