*ICIC Express Letters*
Volume **16**, Number **6**, June **2022**

ICIC International ©2022 ISSN 1881-803X
pp. 613–620

# MULTIPLE LINEAR REGRESSION-SIMULATED ANNEALING-SUPPORT VECTOR REGRESSION TO PREDICT FINANCING ACHIEVEMENT RATE OF AGRI-FOOD CROWDFUNDING

Yaxi Min[1], Liu Lin[1,*], Dan Huang[2], Jingyi Zhang[1], Yuhu Liang[1] and Junhu Ruan[1,*]

[1]College of Economics and Management
Northwest A&F University
No. 3, Taicheng Road, Yangling 712100, P. R. China
minyaxi@nwafu.edu.cn; { zjy00514; yuhuliang_2018 }@163.com
*Corresponding authors: linliu1997@nwafu.edu.cn; rjh@nwsuaf.edu.cn

[2]School of Economics
Dongbei University of Finance and Economics
No. 217, Jianshan Road, Dalian 116025, P. R. China
15663852260@163.com

Received October 2021; accepted December 2021

Abstract. *Although agri-food crowdfunding can solve the difficulties in financing and sales of agricultural products, how to predict the financing achievement rate of agri-food crowdfunding (the ratio of the total amount of the project actually raised by the initiator within the specified time to the pre-set target amount) has not been well resolved. Based on the analysis of related influencing factors, this paper proposes a prediction model MLR-SA-SVR based on the combination of support vector regression (SVR), multiple linear regression (MLR) feature selection and simulated annealing parameter optimization, which is used to predict the financing achievement rate of agri-food crowdfunding. Through the comparisons between models, the results show that the model can achieve higher prediction accuracy in a shorter period of time. According to the prediction model, the initiator can optimize the project settings in advance, predict and increase the financing achievement rate.*

**Keywords:** Agri-food crowdfunding, Financing achievement rate prediction, Multiple linear regression, Simulated annealing-support vector regression

1. **Introduction.** In recent years, crowdfunding has become an emerging method of financing [1]. Crowdfunding includes multiple modes, such as equity crowdfunding [2]. Among them, agri-food crowdfunding belongs to reward-based crowdfunding, which uses agricultural products as a return on investment. Since it can provide products on demand and realize the order production of agricultural products, agri-food crowdfunding has the advantages of alleviating the asymmetry of production and marketing information, reducing circulation links, and reducing costs [3], which is of great significance to agricultural development. Moreover, in China, agri-food crowdfunding contributed to solving the problems of farmers' financing and sales as a former online poverty alleviation model.

There have been many studies on the influencing factors and forecasting of crowdfunding performance. Relatively speaking, the research results on the influencing factors of crowdfunding performance are more abundant (see Section 2 for details). In related prediction studies, most of them predict the success of crowdfunding [4-9]. Among them, Greenberg et al. [5] provided project goals and other characteristics to support vector machine classification algorithms and decision tree models to predict whether crowdfunding

DOI: 10.24507/icicel.16.06.613

can be successful; Wang et al. [7] proposed a deep learning algorithm to predict whether crowdfunding is successful; Yuan et al. [8] designed a framework that can extract latent semantics from the textual description of projects to predict the success of fundraising. Little literature predicted the financing achievement rate as a continuous variable, and there are fewer studies on the prediction of the financing achievement rate of agri-food crowdfunding. Yang et al. [10] used a nonlinear decision model to predict the success of green crowdfunding projects, and converted the success of the project into the ratio of the actual raised amount to the target raised amount; Li et al. [3] used multiple linear regression models to study the factors that prompt people to make rapid investment decisions in agri-food crowdfunding; Li and Du [11] studied the factors that accelerate the achievement of crowdfunding financing goals for the agri-food industry. This paper studies the influencing factors and forecasting methods of the financing achievement rate of agri-food crowdfunding, so as to help project initiator to improve the project setting and increase the financing achievement rate.

In view of this, the main contribution of this paper is to propose the MLR-SA-SVR prediction model for the financing achievement rate of agri-food crowdfunding, and use experiments to verify that the model has better prediction performance. Our research has enriched the performance forecasting methods of agri-food crowdfunding.

The remaining content of this paper is arranged as follows. Section 2 constructs an initial predictive index system; The third section uses regression analysis to select features from the predictive indicators; Section 4 constructs MLR-SA-SVR prediction models, and compares the prediction results between the models; The fifth section summarizes the content of the full text.

2. **Predictive Index System Design.** In order to construct the prediction model of the financing achievement rate of agri-food crowdfunding, we first establish the index system that influences the financing achievement rate. Previous studies mostly focused on the factors that influence the success rate of crowdfunding [3]. Therefore, on this basis, this paper integrates the four main aspects of project information quality, investor perceived value, initiator characteristics, and investor participation to construct an index system that influences the financing achievement rate.

In terms of project information quality, videos [1,12], pictures [11], trademark registration certificates [11], project financing progress [1], social media messages [13], etc. are believed to improve financing performance.

In this paper, minimum investment and target funding are selected to measure the investor perceived value. Li and Du [11] pointed out that increasing the minimum investment can speed up the investment in agri-food crowdfunding. Lagazio and Querci [14] believed that appropriate funding goals can contribute to the success of the project. Some scholars put forward perceptual value such as experience value. For this experience value, this paper selects return level, poverty alleviation story and government support as the measurement indicators. Among them, Kunz et al. [15] believed that the greater the number of rewards, the greater the probability of project success.

In terms of the initiator characteristics, initiators display project and personal information on the project interface to reflect their own characteristics, such as social capital, comprehensive ability, and experience level. Investors will build trust based on these characteristics, thus influencing investment decisions. Shek et al. [16] studied the trust formation mechanism of consumers in online shopping in the form of cases and interviews, and found that it was important for consumers to obtain credit information of the other side, and good credit could reduce uncertainty and enhance consumers' trust. Zheng et al. [17] studied the impact of project initiators' social network capital on crowdfunding project financing performance based on the social capital theory. The empirical results showed that the relationship between project initiators and other project initiators had a significant positive impact on crowdfunding performance.

In terms of investor participation, Mollick [1] pointed out that the greater the number of investors in a project, the greater the possibility of crowdfunding success. Wang et al. [18] pointed out that the number of comments was positively correlated with the success of crowdfunding. Based on the above analysis, this paper builds an initial index system for predicting the financing achievement rate of agri-food crowdfunding. As shown in Figure 1, there are 17 initial indicators in total.

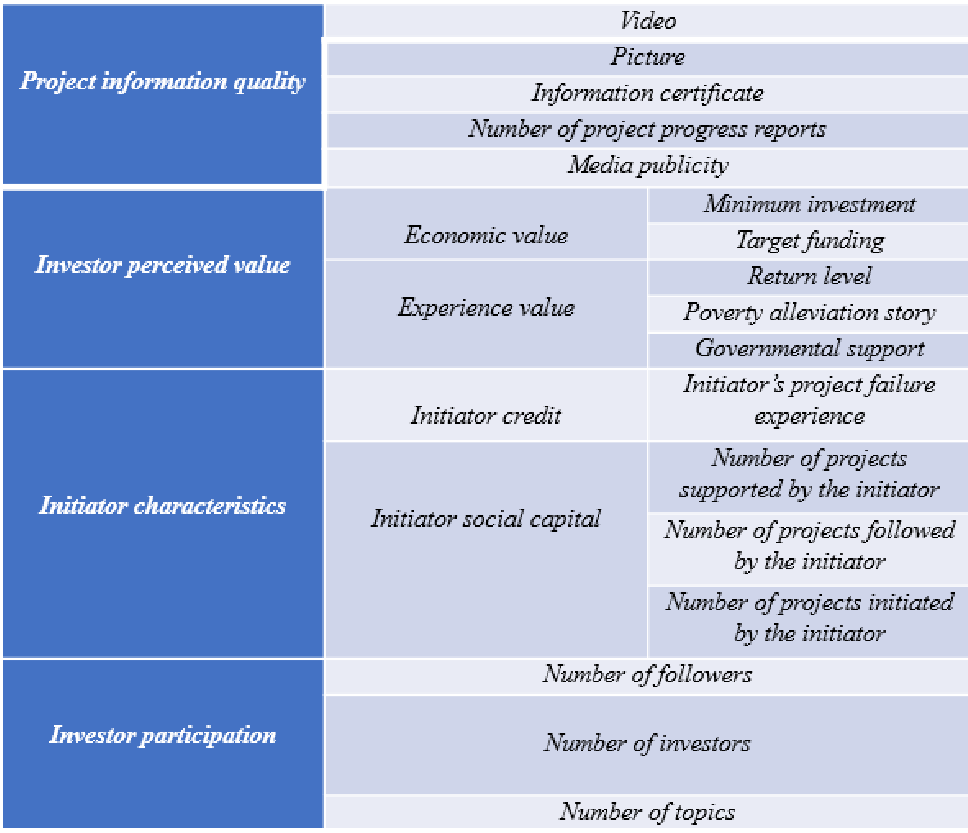


FIGURE 1. Initial index system for financing achievement rate prediction

3. **Feature Selection Based on Regression Analysis.** Ryoba et al. [4] chose 9 features to predict the success of the crowdfunding project, and the results were better than using all the features. Dewi and Chen [19] also pointed out that in machine learning, feature selection is more important than model design. In this work, we select the characteristic indicators that have a significant impact on the financing achievement rate of agri-food crowdfunding based on MLR. On the one hand, we make feature selection for the subsequent machine learning model which is based on regression analysis; on the other hand, we obtained the significant index regression coefficient of MLR in predicting the financing achievement rate.

3.1. **Data sources.** This paper uses a combination method of manual collection and crawler software to continuously collect all successful agricultural crowdfunding projects data from "Jingdong Crowdfunding", the largest product crowdfunding platform in China, from June 2019 to April 2020. At the same time, in order to maintain the stability and randomness of the data, 300 project data were randomly selected as sample data after the elimination of the projects whose target financing amount was more than 100,000 yuan and the financing achievement rate was more than 2,000%.

3.2. **Empirical analysis and results.** The theoretical model formula of MLR is as follows:

$$y = c + \beta_i x_i + \varepsilon \quad i = (1, 2, \ldots, 17) \tag{1}$$

In Formula (1), $y$ represents the dependent variable, namely the financing achievement rate of agri-food crowdfunding; $x_i$ represents each independent variable, namely 17 initial prediction indicators; $c$ is the constant of the equation, $\beta_i$ is the magnitude of the change in the dependent variable when each independent variable changes by one unit, $\varepsilon$ is the random disturbance coefficient of the equation, $i = (1, 2, \ldots, 17)$.

Before the MLR analysis, in order to avoid the problem of multicollinearity, we first performed a correlation test, and the results showed that there was no multicollinearity problem between the two independent variables. In addition, in order to avoid the negative impact of heteroscedasticity on the accuracy and reliability of the regression results, this paper uses the least square method to perform regression on the collected sample data, taking the absolute value of the residual and then taking the reciprocal, the value is then used as the weight for regression.

The results show that there are 10 significant factors influencing the financing achievement rate of agri-food crowdfunding. The number of pictures ($\beta = 2.623$; $p = 0.000$), minimum investment ($\beta = 0.279$; $p = 0.021$), return level ($\beta = 5.627$; $p = 0.029$), number of projects supported by the initiator ($\beta = 0.480$; $p = 0.033$), number of projects initiated by the initiator ($\beta = 20.249$; $p = 0.000$), number of investors ($\beta = 0.259$; $p = 0.000$) and number of topics ($\beta = 1.127$; $p = 0.001$) have a significant positive effect; media publicity ($\beta = -27.140$; $p = 0.021$), target funding ($\beta = -0.007$; $p = 0.000$), and initiator's project failure experience ($\beta = -35.105$; $p = 0.043$) have a significant negative effect.

## 4. MLR-SA-SVR Method and Predicted Results.

This framework mainly predicts the financing achievement rate of agri-food crowdfunding, and proposes a prediction model MLR-SA-SVR based on the combination of SVR, MLR feature selection and SA parameter optimization.

4.1. **MLR-SA-SVR method.** After the feature selection of MLR, a sample containing 10 types of features is generated. Next the sample data is divided into training set and test set. Then, in order to avoid the influence between data of different magnitudes, this paper uses the function mapminmax of Matlab to normalize the data. The data can be converted into a small range of data between $-1$ and 1 through the mapminmax function.

Next, this paper uses SA to optimize the parameters $c$ and $\sigma$ of SVR. As shown in Table 1, and we set a series of initial values of parameters in the SA algorithm. Among them, MarkovLength is the number of iterations at any temperature; DecayScale is the annealing strategy, and we often need the temperature to be cooled in an appropriate way; the initial temperature is the temperature at which the cooling starts.

Table 1. Initial parameters

| Parameter name | Parameter value |
|---|---|
| MarkovLength | 100 |
| DecayScale | 0.85 |
| StepFactor | 0.2 |
| Initial temperature | 8 |
| Minimum temperature | 3 |
| Boltzmann constant | 1 |
| AcceptPoints | 0 |

Then we train the SVR model based on the optimal parameters and the training set data. Finally, we use the training model to make predictions on the test set and output the prediction results. The MLR-SA-SVR model process is shown in Figure 2.

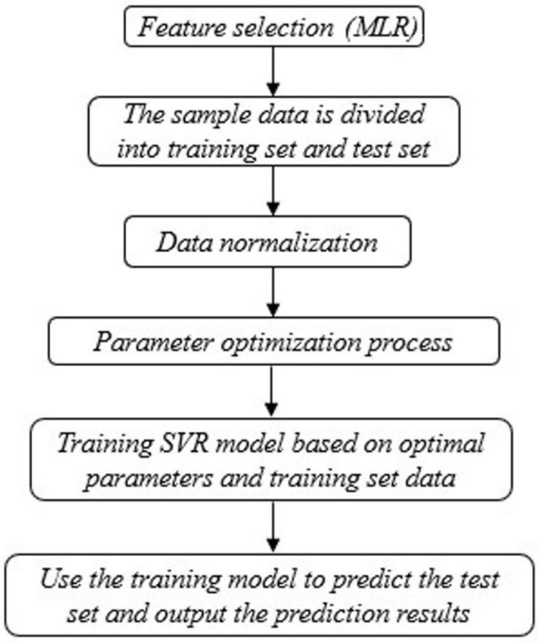


FIGURE 2. MLR-SA-SVR model process

4.2. **Comparison and analysis of SA-SVR and MLR.** First, SA-SVR and MLR are used respectively to predict the financing achievement rate of agri-food crowdfunding. The test sets corresponding to the 5 groups of experiments are respectively 10, 30, 60, 100 and 150 sample data randomly selected from the 300 population samples, and the training set is the remaining sample data from the population samples.

The comparison of prediction errors between SA-SVR and MLR in 5 groups of experiments is shown in Table 2.

TABLE 2. Comparison of prediction results of MLR and SA-SVR

| Number of test data | MAPE | | SMAPE | |
|---|---|---|---|---|
| | MLR | SA-SVR | MLR | SA-SVR |
| 10 | 0.6209 | 0.2363 | 1.3638 | 0.2270 |
| 30 | 0.8515 | 0.2946 | 1.8140 | 0.2674 |
| 60 | 0.9266 | 0.3536 | 1.3224 | 0.3386 |
| 100 | 0.8623 | 0.2346 | 1.6562 | 0.2695 |
| 150 | 0.7204 | 0.2627 | 1.7031 | 0.2560 |
| T-test | 8.743*** | | 13.160*** | |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

As can be seen from Table 2, SA-SVR has a smaller prediction error than MLR. In order to test whether there is significant difference in the mean value of each prediction error between MLR and SA-SVR algorithm in the 5 groups of experiments, T-test was conducted for each prediction error sequence corresponding to MLR and SA-SVR in Table 2. The results show that the $p$-value corresponding to the T-test of MAPE index is 0, and this of the SMAPE index is also 0, indicating that the prediction error of SA-SVR algorithm is significantly smaller than that of MLR. Therefore, compared with MLR, SA-SVR algorithm is more suitable for predicting the financing achievement rate of agri-food crowdfunding, and the prediction accuracy is higher.

In order to further test the stability and correctness of the conclusion that the SA-SVR prediction accuracy is higher than that of the MLR method, this paper conducts a robustness test according to whether there is video in crowdfunding. The results show that, no matter whether crowdfunding projects have videos or not, the prediction errors MAPE and SMAPE of SA-SVR are significantly smaller than those of MLR. It also shows that the relationship between the independent variable and the financing achievement rate is more non-linear.

4.3. **Comparison and analysis of SA-SVR and classic SVR.** SA-SVR, GA (genetic algorithm)-SVR, and PSO (particle swarm optimization)-SVR are all non-linear models, but different algorithms are used to optimize SVR parameters. This section mainly compares the prediction performance of these three models.

In order to avoid occasional problems and better test the performance of the prediction model, this paper randomly divides all 300 sample data into 10 sample sets, each with a total of 30 data, and each sample set serves as a test set. The remaining 9 sample sets are the training set. For each model, we performed 10 times of model training and prediction. Table 3 shows the error values of each model's 10 predictions and their average values.

Table 3. Comparison of prediction results of SA-SVR, GA-SVR, and PSO-SVR

| Serial | T (sec) | | | MAPE | | | SMAPE | | |
|---|---|---|---|---|---|---|---|---|---|
| number | SA-SVR | GA-SVR | PSO-SVR | SA-SVR | GA-SVR | PSO-SVR | SA-SVR | GA-SVR | PSO-SVR |
| 1 | 30.95 | 108.64 | 106.19 | 0.39 | 0.42 | 0.39 | 0.32 | 0.32 | 0.32 |
| 2 | 29.61 | 108.95 | 106.09 | 0.23 | 0.24 | 0.25 | 0.26 | 0.26 | 0.28 |
| 3 | 30.88 | 109.81 | 106.61 | 0.32 | 0.31 | 0.32 | 0.32 | 0.31 | 0.32 |
| 4 | 31.34 | 108.38 | 106.51 | 0.25 | 0.25 | 0.26 | 0.29 | 0.28 | 0.29 |
| 5 | 30.51 | 108.14 | 106.34 | 0.38 | 0.37 | 0.37 | 0.37 | 0.38 | 0.39 |
| 6 | 29.51 | 107.91 | 105.83 | 0.26 | 0.23 | 0.23 | 0.27 | 0.25 | 0.25 |
| 7 | 31.32 | 108.37 | 106.38 | 0.45 | 0.43 | 0.45 | 0.44 | 0.42 | 0.44 |
| 8 | 30.52 | 108.53 | 107.20 | 0.46 | 0.43 | 0.43 | 0.43 | 0.41 | 0.42 |
| 9 | 30.95 | 108.05 | 106.78 | 0.29 | 0.27 | 0.27 | 0.30 | 0.29 | 0.31 |
| 10 | 30.86 | 109.34 | 107.17 | 0.37 | 0.41 | 0.38 | 0.34 | 0.35 | 0.34 |
| Mean | 30.65 | 108.61 | 106.51 | 0.34 | 0.33 | 0.34 | 0.33 | 0.33 | 0.34 |
| T-test | – | −282.61*** | −309.69*** | – | 0.11 | 0.14 | – | 0.26 | −0.07 |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

As can be seen from Table 3, the average running time of SA-SVR was significantly shorter than that of GA-SVR and PSO-SVR, and the error index values predicted by SA-SVR, GA-SVR and PSO-SVR showed no significant difference.

4.4. **Comparison and analysis of MLR-SA-SVR and SA-SVR.** The first two experiments prove that SA-SVR has obvious advantages in prediction accuracy and prediction time compared with linear model (MLR) and classic SVR algorithms, respectively. Then, what are the advantages of the MLR-SA-SVR prediction model based on feature selection in Section 3 compared with SA-SVR that lacks feature selection?

In this paper, all 300 sample data are randomly divided into 10 sample sets, with 30 data in each sample set. Each sample set serves as a test set, while the remaining 9 sample sets are training sets. MLR-SA-SVR and SA-SVR are respectively used to train and predict for 10 times in total. Table 4 shows the error values of 10 predictions of each model and their mean values.

As can be seen from Table 4, the running time of MLR-SA-SVR algorithm is significantly shorter than that of SA-SVR, while the prediction error has no significant difference,

TABLE 4. Comparison of prediction results of MLR-SA-SVR and SA-SVR

| Serial number | T (sec) | | MAPE | | SMAPE | |
|---|---|---|---|---|---|---|
| | MLR-SA-SVR | SA-SVR | MLR-SA-SVR | SA-SVR | MLR-SA-SVR | SA-SVR |
| 1 | 28.33 | 30.95 | 0.36 | 0.39 | 0.32 | 0.32 |
| 2 | 27.77 | 29.61 | 0.23 | 0.23 | 0.26 | 0.26 |
| 3 | 27.31 | 30.88 | 0.32 | 0.32 | 0.30 | 0.32 |
| 4 | 27.87 | 31.34 | 0.29 | 0.25 | 0.32 | 0.29 |
| 5 | 27.59 | 30.51 | 0.40 | 0.38 | 0.40 | 0.37 |
| 6 | 28.23 | 29.51 | 0.26 | 0.26 | 0.26 | 0.27 |
| 7 | 27.33 | 31.32 | 0.36 | 0.45 | 0.37 | 0.44 |
| 8 | 27.84 | 30.52 | 0.45 | 0.46 | 0.42 | 0.43 |
| 9 | 28.21 | 30.95 | 0.29 | 0.29 | 0.29 | 0.30 |
| 10 | 28.32 | 30.86 | 0.40 | 0.37 | 0.35 | 0.34 |
| Mean | 27.88 | 30.95 | 0.33 | 0.34 | 0.33 | 0.33 |
| T-test | −11.76*** | | −0.25 | | −0.19 | |

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

indicating that the prediction accuracy of MLR-SA-SVR can still be guaranteed after MLR feature selection, and the prediction time can be significantly shortened.

So far, the comparison between SA-SVR and MLR proves that SA-SVR is more accurate than MLR, and then SA-SVR is proven to save forecasting time than GA-SVR and PSO-SVR. Last, the comparison between MLR-SA-SVR and SA-SVR proves the superiority of the MLR-SA-SVR in predicting time. Therefore, after these three experiments, MLR-SA-SVR is the best model in this paper.

5. **Conclusions.** This paper mainly studies the forecast of the financing achievement rate of agri-food crowdfunding. First of all, this paper analyzes the factors that influnce the financing achievement rate and uses MLR to perform regression analysis to find the factors that have a significant impact on it. Secondly, this paper proposes the prediction model MLR-SA-SVR based on the combination of SVR model, MLR feature selection and SA parameter optimization, and proves the value of this model in predicting the financing achievement rate of agri-food crowdfunding. This research makes up for the lack of research on agri-food in the field of crowdfunding, and provides more targeted suggestions for the initiators of agri-food crowdfunding projects to increase the financing achievement rate.

**Acknowledgment.** This work is partially supported by National Key Research and Development Program of China under Grant 2020YFD1100601, Shaanxi Science Fund for Distinguished Young Scholars under Grants 2021JC-21 and Tang Scholar of Northwest A&F University.

**REFERENCES**

[1] E. Mollick, The dynamics of crowdfunding: An exploratory study, *Journal of Business Venturing*, vol.29, no.1, pp.1-16, 2014.

[2] G. K. Ahlers, D. Cumming, C. Günther and D. Schweizer, Signaling in equity crowdfunding, *Entrepreneurship Theory and Practice*, vol.39, no.4, pp.955-980, 2015.

[3] Y. Li, J. Du and W. Fu, Thirty days are enough: What determines the crowd's cash time in agri-food crowdfunding?, *China Agricultural Economic Review*, vol.12, no.3, pp.553-575, 2020.

[4] M. J. Ryoba, S. Qu and Y. Zhou, Feature subset selection for predicting the success of crowdfunding project campaigns, *Electronic Markets*, vol.31, no.3, pp.671-684, 2021.

[5] M. D. Greenberg, B. Pardo, K. Hariharan and E. Gerber, Crowdfunding support tools: Predicting success & failure, *Extended Abstracts on Human Factors in Computing Systems (CHI'13)*, pp.1815-1820, 2013.

[6] J. Y. Yeh and C. H. Chen, A machine learning approach to predict the success of crowdfunding fintech project, *Journal of Enterprise Information Management*, 2020.

[7] W. Wang, H. Zheng and Y. J. Wu, Prediction of fundraising outcomes for crowdfunding projects based on deep learning: A multimodel comparative study, *Soft Computing*, vol.24, no.11, pp.8323-8341, 2020.

[8] H. Yuan, R. Y. Lau and W. Xu, The determinants of crowdfunding success: A semantic text analytics approach, *Decision Support Systems*, vol.91, pp.67-76, 2016.

[9] T. Mitra and E. Gilbert, The language that gets people to give: Phrases that predict success on kickstarter, *Proc. of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pp.49-61, 2014.

[10] J. Yang, L. Liu and C. Yin, A non-liner decision model for green crowdfunding project success: Evidence from China, *International Journal of Environmental Research and Public Health*, vol.16, no.2, p.187, 2019.

[11] Y. Li and J. Du, What drives the rapid achievement of a funding target in crowdfunding? Evidence from China, *Agricultural Economics*, vol.66, no.6, pp.269-277, 2020.

[12] Y. Liu, P. Bhattacharya and Z. Jiang, Video-evoked perspective taking on crowdfunding platforms: Impacts on contribution behavior, *The 35th International Conference on Information Systems "Building a Better World Through Information Systems" (ICIS2014)*, 2014.

[13] I. Borst, C. Moser and J. Ferguson, From friendfunding to crowdfunding: Relevance of relationships, social media, and platform activities to crowdfunding performance, *New Media & Society*, vol.20, no.4, pp.1396-1414, 2018.

[14] C. Lagazio and F. Querci, Exploring the multi-sided nature of crowdfunding campaign success, *Journal of Business Research*, vol.90, pp.318-324, 2018.

[15] M. M. Kunz, O. Englisch, J. Beck and U. Bretschneider, Sometimes you win, sometimes you learn – Success factors in reward-based crowdfunding, *Multikonferenz Wirtschaftsinformatik*, pp.467-478, 2016.

[16] S. P. Shek, C. L. Sia and K. H. Lim, A preliminary assessment of different trust formation models: The effect of third party endorsements on online shopping, *The 36th Annual Hawaii International Conference on System Sciences*, 2003.

[17] H. Zheng, D. Li, J. Wu and Y. Xu, The role of multidimensional social capital in crowdfunding: A comparative study in China and US, *Information & Management*, vol.51, no.4, pp.488-496, 2014.

[18] N. Wang, Q. Li, H. Liang, T. Ye and S. Ge, Understanding the importance of interaction between creators and backers in crowdfunding success, *Electronic Commerce Research and Applications*, vol.27, pp.106-117, 2018.

[19] C. Dewi and R.-C. Chen, Random forest and support vector machine on features selection for regression analysis, *International Journal of Innovative Computing, Information and Control*, vol.15, no.6, pp.2027-2037, 2019.