

How to Succeed in Crowdfunding: Setting a perfect goal

Thanh Tran
Utah State University
Logan, UT 84341
thanhtd.ithut@gmail.com

Hongkyu Choi
Utah State University
Logan, UT 84341
hongkyu.choi@aggiemail.usu.edu

ABSTRACT

To find the answer of how to succeed in crowdfunding, we find many previous works which are mainly focused on to recommend appropriate investors to a project or investigate more influential features on its success. However, in this paper, we would like to suggest a way to set a perfect goal with introducing a regression model. Establishing a goal appropriate is essential for project to success eventually as well as predicting the number of backers a certain Kickstarter project.

Our proposed model show better performance than existing algorithm(Simple linear regression and XgBoost). It returns RMSE(Root Mean Square Error) 23.7 and 1,659, prediction for the number of backer and funding respectively.

1. INTRODUCTION

Why is this an interesting question to ask and why would we care about the answer to this question or a solution to set a perfect goal? The Kickstarter policy is one or nothing. It means that if the amount of final pledged money is equal or higher than the goal, the fundraising campaign is successful. Otherwise, the fundraising campaign is fail and the creator receives nothing. If the creator overestimate the project goal, they will be fail, but if the creator set the project goal too low, they may not get attraction from community and gain less amount of money. Hence, predicting the amount of money or the number of backers are very important for Kickstarter creators to raise fund.

To our knowledge, There is only one research work try to predict how much fund a Kickstarter project can receive. Particularly, researchers at [4] convert the problem of predicting the amount of final pledged money into classification problem by dividing such amount of final pledged money into different range. So the problem of predicting how much fund a Kickstarter project can receive will be the problem of predicting which range of money the project can receive.

Comparing to this work, our work is totally different, instead of predicting the range of final pledged fund, we build

regression model to predict exactly the amount of money as well as the number of backers given a certain Kickstarter project.

Our contributions (or research questions) in this proposal are:

- Understand the influence of multiple factors toward the number of backers and the amount of final pledged money that a certain project can receive. We will show statistic values to illustrate for such influence.
- Given a project, we build a model to predict how much pledged fund the creator can receive
- Building a model to predict how many backers will fund for the project.

2. RELATED WORKS

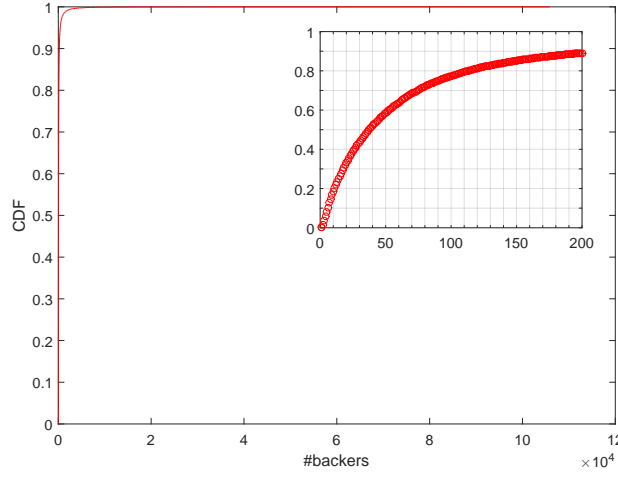
There exists many research works on crowdfunding problem. We explain previous works based on trend of crowdfunding research.

Researchers have predicted whether the project can be successfully funded or fail. [6] collected 13,000 projects on Kickstarter and extracted 13 features from each one to develop a classifier to predict project success with 68% accuracy. [5] extends the work and show how the temporal amount of money can help improve the accuracy. [8] focused on text features of project pages and show how using phrases features to predict project success.

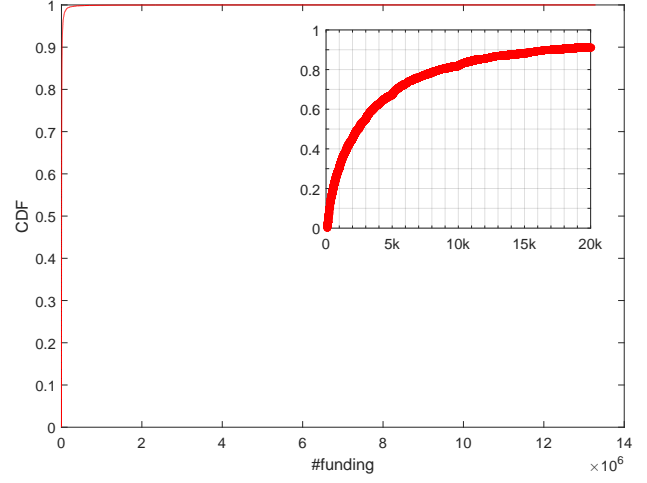
Another research trend tries to correlate social media activities during running fund raising campaign to project success and proposed solutions for investor recommendation problem. [7] studied how the amount of money can be affected by promotional activities on social media like Twitter. [1] used promoter network on Twitter to show the correlation between the connectivity of project promoters and project success. They also developed backer recommendation in which potential investors are suggested. [3] proposed different ways of recommending investors by using hypothesis-driven analysis of pledging behavior. [2] presented various factor influenced investor retention which allows to identify different groups of investors.

Comparing with the previous research work, we collected largest dataset consisting of more than 150k projects. Our problem is totally different comparing to existed works. That is, we construct statistical models that examine multiple predictive factors toward building two models: (i) one predicts the number of backers will back for the project and (ii) the another predicts the amount of pledged money that the project can receive.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.



(a) CDF of Backers Distribution



(b) CDF of Funding Distribution

Figure 1: Cumulative Distribution Function of Backers and Funding

3. ANALYSIS

In statistics, linear regression is an approach for modeling the relationship between a dependent variable y and one or more explanatory variables denoted X . In this paper, there are more than one explanatory variable, the process is called multiple linear regression.

In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models.

As Figure 1 shown, only 10% of total project obtains more than 200 backers and \$20,000 funding. In other words, majority of project could not achieve such a large number of backers and funding. It confirms that it is essential to success to know how many backers will fund and how much money will be collected.

Before discovering a model, we would like to see relation between major features and the number of backers, finally pledged money. In Figure 2, it gives that there is not specific linear relation between independent variables such as project duration, description length, the number of rewards, and the number of frequently asked question.

It is likely to have a shape of normal distribution in that there are more backers and participation in fund with middle of independent variables. In case of duration, there are many marks near point of "30". It has similar pattern in relation on the number of reward. This phenomenon makes hard to make model because the different dependent variables from the same independent variable. It requires to find some transformation method to fit a model.

There are similar patterns of distribution whether response variable is final pledged money or the number of backers. To understand this phenomenon, it is need to investigate relationship between them. It is obvious that there should be positive relation between the number of backers and pledged money. When more users are involved in a specific project, it is natural that the project have more fund in the last.

Figure 2 shows relation between the number of backers and final pledged money. Linear regression line is represented with the below equation.

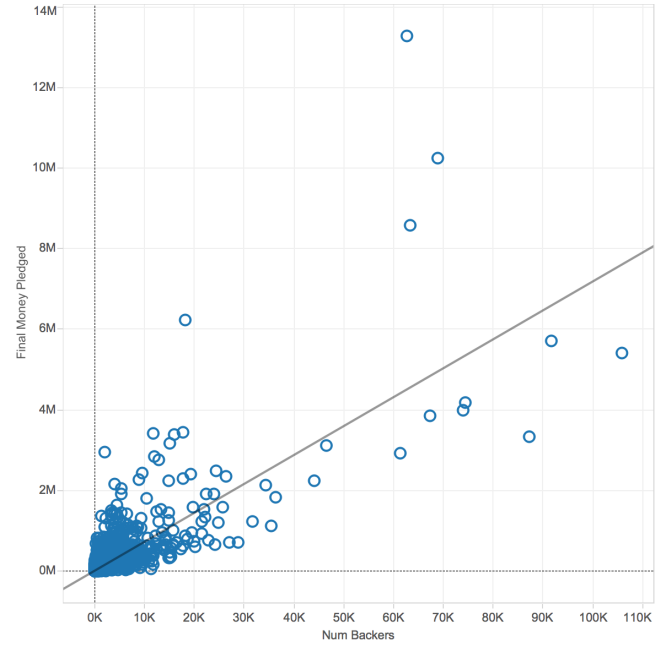


Figure 3: Scatter plot between Final pledged money and backer

$$FinalMoneyPledged = 71.8998 * NumBackers + 845.879 \quad (1)$$

With significance level 0.05, the model is acceptable because its p-value is less than 0.0001. Besides, R^2 which provides a measure of how well future outcomes are likely to be predicted by the regression model and how well the regression line fits real data is 0.6321. It represents that this model with the number of backers explains 63.21% variance of final pledged money.

Figure 4 shows relation between final pledged money and the number of backer with previous project success rate. It

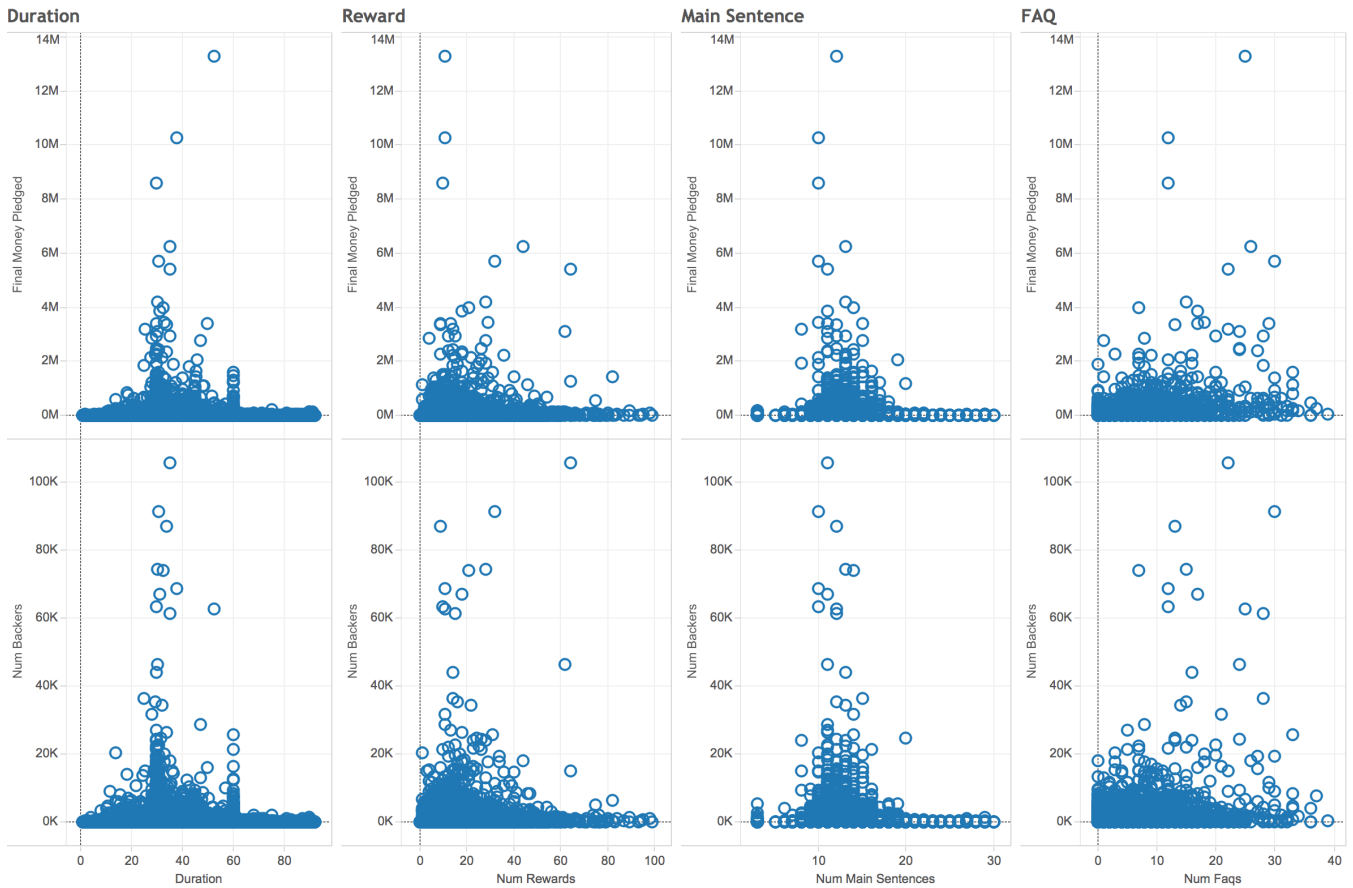


Figure 2: Scatter plot between available features and Final pledged money, backer

represents that the more success rate in previous project, the more chance to attract backers which leads to get more financial support. This could not fully guarantee successful modeling but it point out one of features which affect project success.

4. EXPERIMENTS AND RESULT

5. CONCLUSION

6. FUTURE WORKS

CCS Concepts

•Computer systems organization → Embedded systems; Redundancy; Robotics; •Networks → Network reliability;

Keywords

ACM proceedings; L^AT_EX; text tagging

7. REFERENCES

- [1] *What Motivates People to Invest in Crowdfunding Projects? Recommendation using Heterogeneous Traits in Kickstarter*, 2015. Rakesh, Vineeth and Choo, Jaegul and Reddy, Chandan K.
- [2] T. Althoff and J. Leskovec. Donor retention in online crowdfunding communities: A case study of donorschoose. org. In *Proceedings of the 24th International Conference on World Wide Web*, pages 34–44. International World Wide Web Conferences Steering Committee, 2015.
- [3] J. An, D. Quercia, and J. Crowcroft. Recommending investors for crowdfunding projects. In *Proceedings of the 23rd international conference on World wide web*, pages 261–270. ACM, 2014.
- [4] J. Chung and K. Lee. A long-term study of a crowdfunding platform: Predicting project success and fundraising amount. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, pages 211–220. ACM, 2015.
- [5] V. Etter, M. Grossglauser, and P. Thiran. Launch hard or go home!: Predicting the success of kickstarter campaigns. In *COSN*, 2013.
- [6] M. D. Greenberg. Public online failure with crowdfunding. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pages 333–334. ACM, 2015.
- [7] C.-T. Lu, S. Xie, X. Kong, and P. S. Yu. Inferring the impacts of social media on crowdfunding. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 573–582. ACM, 2014.
- [8] T. Mitra and E. Gilbert. The language that gets people



Figure 4: Scatter plot between Final pledged money and backer with previous success project rate

to give: Phrases that predict success on kickstarter. In *CSCW*, 2014.