

Project 1

With the rapid development of information technology, crowdfunding (one sub-field of Internet Finance) has become a new and creative channel for both individuals and small-business owners to raising money to support their dreams. In this exciting project, you are required to use data mining technology to explore a provided crowdfunding data. I collected this real data from Kickstarter Platform by using a web crawler. It contains 3, 5683 real projects posted on that platform range from 2009/4/24 to 2012/7/29. For each project, we recorded its name, url-link, description, # backers and many other features (see the column of the provided data. Most of the provided features are self explanatory and their detailed explanations can be found in the second sheet of our csv file). **You are encouraged to:**

- (1) Exploring the data to show the general fundraising pattern of project in different category;
- (2) In your opinion, what factors can influence the fundraising performances of a project (i.e., fundraising amount, fundraising results) on that platform? Can you verify them one by one through the provided data?
- (3) Can you build up some models to predict the status of a crowdfunding project (i.e., success or not)?
- (4) Can you build up some models to predict the fundraising amount of a crowdfunding project (i.e., regression model) on that platform?
- (5) Based on your analysis, can you give some valuable suggestions to fundraisers on the platform to help them raise more money for their projects?
- (6) Any good suggestions to investors to help them select project with high quality on that platform?

I hope to see great projects from your guys.

Notes:

- (1) You just need to submit a report in a word version. The code you used in this project can be listed in the appendix part of your report.**
- (2) Please do not submit a R script directly to me !!!!!**
- (3) The project deadline is 11/26/2018**