

Elizaveta Kadetova
Moscow Higher School of Economics, 2019-2020
Research project on Computational linguistics
Proposal

Crowdfunding for charity: correlation between the features of a campaign description and its success.

Purpose of the Study

The purpose of the study is to find out whether there is any correlation between particular characteristics of a text aimed to raise money for some charity cause and the success of the collection. It is important to note here that by correlation we should mean just regularity, not causality. For the purpose of the study materials of charitable campaigns from a popular Russian crowdfunding platform will be used.

Terms

Campaign - a project aimed to collect a particular sum of money for a particular purpose within a limited period of time.

Charitable campaign - a campaign aimed to help people or animals, or to solve some social problem.

Founder - a person or an organisation who/which launches and owns a campaign.

Campaign description - a text published by a founder of a campaign and providing information about its goal.

Donations - sums of money given without a reward to support a (charitable) campaign.

Background

Crowdfunding is a way to attract money for financing a project. It is called that way because the idea is to gather relatively small amounts of money from many people, so that the total amount makes a project possible. Although it has become well known and popular recently due to the Internet and electronic payments services, it is not an invention of the 21st century. For example, in 1850-80s the National Theatre in Prague was built and later restored after a fire due to common people's donations. And this is not the only, and not the earliest, case. Interestingly, there is an inscription inside saying "*Národ sobě*", which means "people for themselves", as a reminder that this is a result of the collective impact of those who wanted to have a theater in the city. It is somewhat close to what modern fundraisers frequently offer to their backers: a mentioning of their names in the titles of the movie or on the cover of the album, released due to their funding.

Modern crowdfunding on the Internet began to develop around 2000s. Today there are several clearly distinguished models of raising funds online, such as reward-based, debt-based, donations-based, equity-based, and others. The causes are various: from creating and launching a new product into market to paying off one's penalty. Naturally, charity is among the most popular causes for crowdfunding.

Crowdfunding seems to be constructed rather simply, but it is not. Every campaign profile consists of many details. Some campaigns are successful, others are not, and those successful ones show different rate of luck. But there is no answer why it is so. Definitely it is a combination of factors: the idea behind or the purpose of the collection, target amount of money, reputation of the founder, visual design of the campaign, for example, presence of high-quality photographs or a promo video, rewards and gifts, advertising, and so on. This is all applicable to charity (donations-based crowdfunding) as well. However, there are some points that make charitable crowdfunding specific.

Relevance

Donations-based charitable campaigns, to my opinion, are more, let's say, vulnerable. Unlike reward or equity-based campaigns they don't offer any gain to donors, except for small thank-you souvenirs like a magnet, a bracelet, or a ticket to a partnership cinema. And they don't have any product idea behind. So charitable campaigns suggest that users just give their money away. That is why the role of the text, convincing them to do so, increases as it becomes the main tool for a founder.

According to my experience of working for different charities, fundraising texts are indeed among the most controversial factors nonprofits' staff constantly argue about. Should they rather be brief and clear not to bore a reader, or vice versa, detailed in order to be credible? Is it good to use specialised, for example medical, terms? Does it make sense to include direct speech? What style is preferable: professionally calm or emotionally expressive? There are many opinions on this issues. Really talented authors create texts everyone likes, but no one knows why. Why does one text, as they say, "work", and another one does not?

An answer to this question hardly lies in just measurable features. However, I believe some interesting patterns could be revealed there, which may lead to further assumptions and be useful for nonprofits in terms of increasing fundraising effectiveness, not only via special platforms, but in general.

Studies of the Topic

In recent years, crowdfunding success factors have been the subject of interest for many researchers from different countries. There are many popular articles and research papers devoted to this issue. Several large datasets are available on Kaggle, where users suggest

models predicting if a project is going to be successful. Success factors of Russian crowdfunding projects, based on data from Boomstarter platform, were explored as a subject of master's thesis by a student of St. Petersburg University Graduate School of Management Anna Petrova, which is claimed to be a pioneered paper in studying success factors of Russian crowdfunding projects.

However, I am not aware of any studies 1. devoted specifically to charitable campaigns 2. focused on texts.

Data Sample

The vast majority of large and medium Russian nonprofits are engaged in public fundraising on the Internet using their websites, mass media, social media, and crowdfunding platforms. For research purposes, campaign profiles from crowdfunding platforms are the best match since multiple data are publically available and well organised there.

Russian nonprofit organisations most often use the following platforms designed for public fundraising: Boomstarter.ru, Planeta.ru, Dobro.mail.ru, and an information portal Takie Dela. The latter is very popular and effective, but is not of much help for data study since the main tools there are not projects' profiles per se, but professional reports and picture stories, provided by the edition, with no public evidence of the impact of each particular piece to the progress of a campaign.

It is also worth mentioning that all platforms have their own policies and technologies regarding campaign promotion, termination or abortion, which frame its progress and result, so it seems appropriate to analyse the platforms separately first.

The data available on the three above mentioned platforms are not fully the same.

Planeta.ru: title, short description, full description, target goal, goal achieved, number of transactions, **date of closure, date of start.**

Boomstarter.ru: title, short description, full description, target goal, goal achieved, number of transactions, **date of closure.**

Dobro.mail.ru: title, short description, full description, target goal, goal achieved, number of transactions.

So only Planeta.ru allows to see how long it took to achieve the goal, which is important as it can be considered as one of the indicators for measuring success. Taking this into account, I have chosen this platform as the source of data for this project.

Success Evaluation

Classic crowdfunding model suggests that a goal or at least a particular benchmark must be reached within originally specified period of time, otherwise all the pledges are canceled. Charitable campaigns on the platforms mentioned above are based on co-called flexible funding principle. A campaign is considered successful and the founder can get the money no matter what part of the target sum is gathered. And if the target sum is gathered, a project is not closed automatically, but goes on until its founder closes it. Often these campaigns terminate with a result of 150-200% of the original goal. Thus, a simple classification of such campaigns into successful and unsuccessful is impossible, and more complicated approach should be figured out, taking into account not only the total amount of money collected, but the number of donations (if it was mostly covered by a big donation from one generous donor, it is not a typical situation for crowdfunding), the duration of the collection, and the target goal ratio.

Research Design

In order to achieve the goal of the study the following objectives can be set:

1. To collect and systematise the data needed. The dataset is expected to consist of 2000 texts (or text id-s) of campaigns with their indicators: target goal, goal achieved, number of transactions, date of closure, date of start.
2. To assess the data and set an approach (index) of marking a project as successful or not successful (possibly: more or less successful).
3. To create a corpus of descriptions and fulfil a reasonable text preprocessing.
4. To analyse texts for a number of characteristics (see below)
5. To match significantly different characteristics with the success indicators.
6. To find patterns and make assumptions.

Linguistic characteristics of the texts, which can be analysed, include:

1. Text length in characters
(Premise: it is believed that the shorter the text, the more people will read it, as people mostly don't like to read a lot)
2. An average length of the sentence in each text
(Premise: it is believed that long sentences make a text harder to perceive and thus reduce the amount of people eager to deal with it)
3. Presence of numbers in the text
(Premise: it is believed that detailed descriptions, which include prices, dates, ages, etc., are more trustful)
4. Presence of words spelled in upper case, except abbreviations

- (Premise: some authors believe that spelling whole words in capitals (e.g. PLEASE! NEED HELP!) attract attention and increase involvement, others say it is bad form, which can cause nothing but irritation)
5. The most frequent lexemes
(No particular premise so far)
 6. Frequency of diminutives
(Premise: it is believed that diminutive forms of words emotionally affect readers and cause sympathy)
 7. Named-entity recognition
(Premise: it is believed that real names of people or places make the story more trustful and help readers identify themselves with those who need help)
 8. POS distribution of the most frequent words
(No particular premise so far)
 9. Presence of direct speech (quotation)
(Premise: it is believed that direct speech (e.g. words of a nonprofit's director or a mother asking to help her child) makes the story more vivid and brings more involvement)
 10. Possibly: presence of emotionally charged vocabulary
(Premise: some authors believe that very expressive texts increase involvement, others think it looks unprofessional and kills trust)
 11. Possibly: topic modeling
(Premise: it is believed that some causes are more likely to be successful than others, e.g. people would rather support a child than a senior or a homeless person, but rather a homeless person than an infrastructure project, for example, a regular work of a hotline for orphan diseases or an educational conference).

The project is planned to be carried out in Python 3 in Google Colab.

Further Development

In future the study can be expanded to cover:

1. Other Russian-language platforms, listed above, to verify the patterns found within current research of the Planeta.ru campaigns
2. Crowdfunding texts in Czech, Polish, and English, as it has been already revealed that the same features of campaigns may be involved in completely different patterns depending on a country and audience¹.

¹“In favor of this argument, Zheng *et al.* (2014) have found that there is a negative relation between campaign's duration and its success on the United States crowdfunding market, while positive one on Chinese crowdfunding market.” [Petrova A. 2018. Success factors of Russian crowdfunding projects: empirical study of Boomstarter.ru platform.]