Abstract

Unfinished

1. INTRODUCTION
2. Background

According to **Crowdfunding in a Canadian Context launched by Canada Media Fund (CMF) in 2012**, Crowdfunding is the raising of funds through the collection of small contributions from the general public using the Internet and social media. There are some benefits of crowdfunding. First, it promotes innovation in society, because crowdfunding is used to raise money to fund the development of a singular project or a project in the development stage. Additionally, acting as a social media platform, some crowdfunding platforms are more accessible for the public. They facilitate a community where individuals can fund their interests and passions. Second, crowdfunding lowers a barrier to finance a project. Instead of using conventional method, like getting a loan from a bank or issuing stock, a corporate has increased its access to raise money through crowdfunding. Conventionally, a bank is usually unwilling to credit money to an entrepreneur, since a startup company has unstable cash flow, and it is difficulty for a bank to evaluate its business risks. On the other hand, regulated by the listing rule and having a limited budget to pay listing fee, a startup company also has difficulty raising money through issuing stock **(Crowdfunding in a Canadian Context, Canada Media Fund (CMF), 2012)**. Third, crowdfunding generates publicity and attention for a project. A crowdfunding network can be a project’s best cost-free promotion, because it allows a project creator to engage more closely and more directly with investors or potential consumers through direct connection and deeper engagement with the development and production process of a product **(The Economist. Sept 02, 2012)**.Forth, crowdfunding provides many creative content developers with a mean of concept validation or beta testing. Through crowdfunding campaigns before launching a product, a creator can collect vital information from public reactions and critiques to a project’s concepts and features. The feedback can be helpful for business and editorial decision making that respond directly to potential customers’ needs and desires **(Crowdfunding in a Canadian Context, Canada Media Fund (CMF), 2012)**.

1. Brief History

Crowdfunding got attention in the United States when ArtistShare was launched in 2003. It is a fundraising platform for film/video and photography projects as well as music, in which an artist can seek donations from fans **(Freedman, D. M., & Nutting, M. R., 2015).** Until now, projects on ArtistShare have received six Grammy awards and fifteen award nomination **(Crowd Funding: How to Raise Money and Make Money in the Crowd, Modwenna Rees-Mogg)**. After ArtistShare’s success, more rewards-based crowdfunding platforms appeared. Two of the most prominent platforms were Indiegogo in 2008 and Kickstarter in 2009. In addition to the arts, they also provide funding campaigns for social causes, entrepreneurs and small businesses **(Freedman, D. M., & Nutting, M. R., 2015)**. Until now, there are over $4.6 billion (USD) have been pledged from over 17 million backers in Kickstarter. Additionally, there are 174,593 projects have been successfully funded **(Kickstarter)**. Among projects in Kickstarter, Pebble Watch is one of the most successful crowdfunding campaigns. It is a smart watch made by the Pebble Technology Company. In less than a week the number of the raised money had surpassed the previous Kickstarter record of $3.3 million (USD). Finally, Pebble Technology had reached $10 million (USD) which was hundred times of its funding target **(Crowdfunding in a Canadian Context, Canada Media Fund (CMF), 2012)**. Since 2009, crowdfunding has emerged as a major funding source. From 2009 to 2015, as the rise of Kickstarter and IndieGoGo which caused an explosion of crowdfunding platforms, there arose an urge for crowdfuning. Crowdfunding revenue rose from $530 million in 2009 to $1.5 billion in 2011. By 2012, there were more than 450 crowdfunding platforms, which raised more than $2.7 billion in the world. By 2015, that number had grown to $24.4 billion **(startups.com; A Brief History of Crowdfunding)**. Furthermore, in 2012, crowdfunding gained Washington’s support in the USA. President Barack Obama signed the Jumpstart Our Business Startups (JOBS) Act into law, which reduced regulation on small businesses and legalized equity crowdfunding **(Crowdfunding in a Canadian Context, Canada Media Fund (CMF), 2012)**.

1. Structure of Crowdfunding

There are three roles participating in a crowdfunding campaign: platforms, creators, and backers. First, a platform serves as an intermediary, which matches fundraisers and funders to each other. Second, a creator can be a fundraiser, entrepreneurs and others who are intended to raise fund from supporters in a crowdfunding platform. Third, a backer can be an investor, contributor or funder, also defined as the ‘crowd’ from the term of crowdfunding. A backer financially supports a project a creator launches on a platform and expects the return on money invested. In order to operate business, a platform generally charges fundraisers a fee, which based on the amount of money raised. For example, Kickstarter charges fundraisers a flat 5% fee from the project’s funding. Indiegogo, charges fundraisers a 4% fee from the project’s funding if a certain fundraising goal is met or 9% if a fundraising goal is not met **(Tomczak, A., & Brem, A., 2013)**.

There are four categories of crowdfunding: donation-based, reward-based, lending-based, and equity-based. First, in donation-based crowdfunding, a backer contributes money to a project with no compensation**.** Second, in reward-based crowdfunding, a backer invests in a project in exchange for non-financial rewards **(Giudici, G., Nava, R., Rossi Lamastra, C., & Verecondo, C., 2012)**. A Reward-based crowdfunding campaign is commonly categorized into two types of financing models:“Keep-It-All” (KIA) and “All-Or-Nothing” (AON). In the KIA model, a creator keeps the entire amount of money raised whether or not the capital raising goal is met. A backer thereby needs to bear risk that the amount invested cannot be withdrawn from a creator when the goal is not achieved. In the AON model, by contrast, a creator keeps nothing unless the goal is met, thereby shifting risk to backers **(Douglas, J., Cumming, Gael, Leboeuf, Armin, Schwienbacher, 2015).** The most two prominent reward-based crowdfunding platforms are Indiegogo and Kickstarter. AON model is the main part of Kickstarter while Indiegogo allows both AON and KIA campaigns **(Freedman, D. M., & Nutting, M. R., 2015).** Third, in lending-based crowdfunding, a backer lends money to a project with the expectation that contribution can be reimbursed, perhaps with interest **(Giudici, G., Nava, R., Rossi Lamastra, C., & Verecondo, C. , 2012)**. There were $522 million (USD) being raised globally in 2011 through ending-based crowdfunding. In terms of the amount of money raised, lending-based crowdfunding became the second-largest form of crowdfunding **(Crowdfunding in a Canadian Context, Canada Media Fund (CMF), 2012)**. Forth, in equity-based crowdfunding, a backer receives equity so as to buy ownership in a creator’s company or right in a project or earns a share of the revenue or profits in a project. A 2012 study shows that just 15% of crowdfunded projects in 2011 were equity-based, while 43% were reward-based, 28% were donation-based, and 14% were lending-based **(Bannerman, 2013).**

1. Motivation and Purpose

Crowdfunding tends to be more successful and become a popular method for investment, seed money and start-up funding **(Tomczak, A., & Brem, A., 2013)**. Not only the total amount pledged but also the number of projects has significant growth. Take Kickstarter for instance, one of the world largest crowdfunding platforms. The number of projects in 2015 was over 80 times as much as in 2009; the total amount pledged had also grown from 1.6 million to more than 600 million US dollars in the same period **(icopartners.com; Kickstarter in 2015–Review in numbers, 2016)**. Moreover, crowdfunding also provides entrepreneurs with a useful alternative way to raise funds when they face difficulty attracting financial supports from bank loans or equity **(Belleflamme, P., Lambert, T., & Schwienbacher, A., 2010)**.

Despite the popularity of crowdfunding, however, many projects could not reach their funding goals within the prescribed period. The statistics shows that there are only one-third of projects successfully funded on average. In 2015, only 27.19% of projects met their goals **(icopartners.com; Kickstarter in 2017 – Year in review, 2018)**. In terms of the low success rates in the project, even if crowdfunding benefits entrepreneurs with a more efficient method to collect start-up funding and provides an investment opportunity for investors, there are also some risks existing for fundraisers and investors. As a result, if the risk in crowdfunding can be managed, the success rate in the project will be improved likely so as to benefit the participant in crowdfunding.

In order to manage the risk and enhance the success rate in the project, this paper aims to reduce information asymmetries in transaction. Given that crowdfunding is mostly conducted through e-commerce channels, customers gain inadequate physical information about a product, lack face-to-face communication with sellers and miss possibility to evaluate a product’s quality prior to purchase. Additionally, since most of the projects are innovative and haven’t exited before, no recommendations backers can take prior to invest. Thus, backers only can rely on information the project initiators provide. Having restricted information and avoiding making mistakes, customers tend to refuse to purchase under uncertainty **(Gierczak, M. M.; Bretschneider, U. & Leimeister, J. M., 2014)**. As a result, in order to improve the success rate in the project, it is necessary to reduce information asymmetries between fundraisers and investors so as to increase buyer motivation.

In terms of the solution to reduce information asymmetries, this paper aims to evaluate the success rate and the final pledge amount of a launched project at the early stage. Through evaluating the success rate of a project at the early stage, backers can foresee which projects have higher probability to succeed and avoid wasting their money on a project with a highly failed rate. On the other hand, fundraisers also can examine a project with a highly failed rate and improve its quality as soon as possible. Moreover, with the predicted final pledge amount of a project, fundraisers will have more insight into the value of their projects. Fundraisers will know how much the predicted pledge amount higher than the funding goal they set. If the predicted pledge amount of a project is much higher than the funding goal, a project is likely popular with the market, and it is worthy for fundraisers to put resources into project development. On the other hand, if the predicted pledge amount of a project is near the funding goal, fundraisers will still have some rooms to improve their projects.

Stacking, one of ensemble methods in machine learning, is employed to predict the success rate and the final pledge amount of a launched project. Ensemble learning techniques aim to combine the prediction of the different classifiers to generate a final hypothesis, typically by voting, averaging or using a meta-level-classifier **(Sikora, Riyaz and Al-laymoun, O'la Hmoud, 2014)**. Ensemble methods are usually used to overcome three types of problems which are suffered from base learning algorithms: the statistical problem; the computational problem; and the representational problem **(Thomas G. Dietterich, 2000).** Experimental evidence in Freund and Schapire (1996) and other studies summarized in Dietterich (1997) have shown that ensemble methods are often much more accurate than any base learning algorithms **(M.A. Arbib, Ed., 2002)**.

1. LITERATURE REVIEW

Unfinished

1. METHODOLOGY
2. Data Collection

The research target is the project on Kickstarter and Indeigogo, both of which are global fundraising platforms. Since they are the two most popular donation-based crowdfunding platforms, the data collected for the research is abundant and the source of data is also more trustworthy. The dataset is collected from Kaggle, which is an online community of data scientists and learners, owned by Google. It offers machine learning competition and allows user to share the dataset. The dataset is constituted of historical projects from Kickstarter and Indeigogo and includes 593011 records from May 2009 to March 2018.

The records respectively from Kickstarter and Indeigogo occupy 55.89% and 44.11% of the whole dataset. The number of successful projects is recorded as 213611, accounting for 36.02% of total projects while the number of failed projects is 379400, around 63.98% of total projects. The average amount of pledge is 11900 US dollars, the average amount of goal is 31500 US dollars, and the average number of backers is 125. There is also a large gap in the amount of pledge, the amount of goal and the number of backers among projects, since there is high standard deviation and large gap between maximum and minimum in these three columns. (TABLE II) There are six columns for each project in the dataset, which are status, category, location, backers, goal\_usd and pledge\_usd. The column of category consist 46 categories. The top 10 categories which have ten of the most numbers of projects are: Music, Film & Video, Film, Technology, Publishing, Art, Design, Games, Fashion, and Food. The project currency is mainly based on USD dollars. There are 80.75% of projects in the dataset, whose currency is based on USD dollars. TABLE I describes detailed data information, including column name, data type and short descriptions. Another observation in descriptive statistics is given in TABLE II.

There are two datasets offered by Kaggle: Indeigogo and Kickstar. In order to maintain the data’s diversity and quantity, both datasets are merged into a whole dataset. Only a column which appears in both Indeigogo and Kickstar dataset is chosen into the new dataset. Therefore, the columns displayed in the TABLE I are common ones in the both dataset. Furthermore, before the merge of dataset, the records which have missing values respectively in the Indeigogo and Kickstar dataset are removed as data leakage.

TABLE I. DATA TYPE AND DESCRIPTION OF EACH COLUMN



TABLE II. DESCRIPTIVE STATISTICS OF DATA COLUMNS



1. Exploratory Data Analysis

Exploratory Data Analysis (EDA) which is a method to summarize the characteristics of data with graphical and numerical methods functions to give more insight into data.

Figure 1 displays the number of successful and failed projects by different categories. The number of failed projects is larger than the number of successful ones in most categories. Only in categories of Comics and Theaters does the number of successful projects exceed the numbers of failed projects. On the other hand, the total number of successful projects accounts for 36.02% of total projects. There are fourteen categories in which a percentage of successful projects are higher than this number. They are Comics, Theater, Audio, Tabletop\_games, Music, Games, Film & Video, Gaming, Art, Design, Camera\_gear, Dance, Politics and Home.

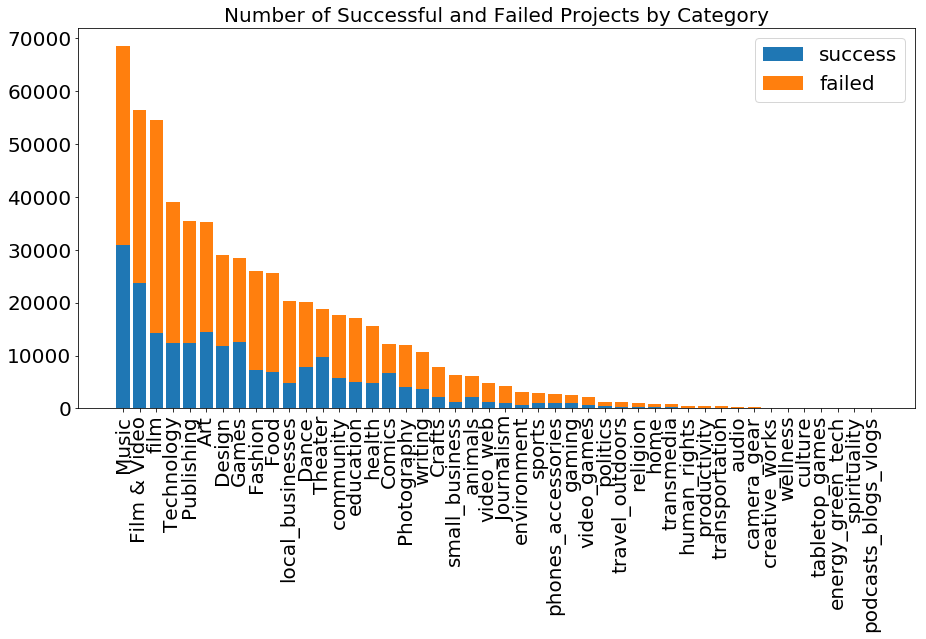


Fig 1. Number of Successful and Failed Projects by Category

Among all the categories, a category of Energy\_green\_tech has the highest average amount of pledge. Under the top 10 categories which have ten of the most number of projects, a category of technology has the highest average amount of pledge. A project which has the most amount of pledge, 20338986 US dollars, is located in a category of Design. (Figure 2)

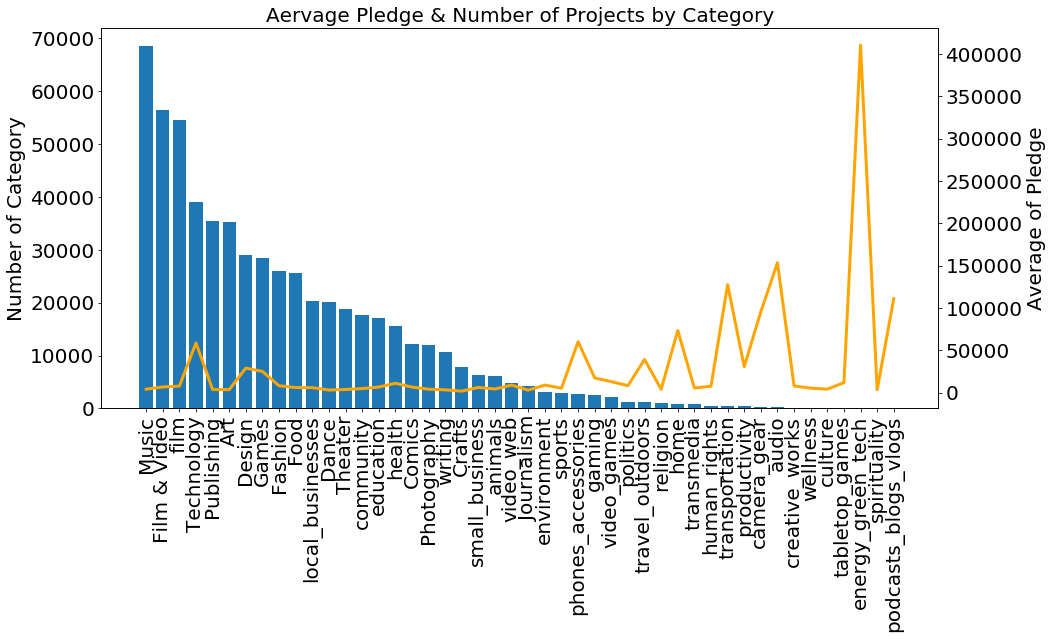


Fig 2. Average Pledge & Number of Projects by Category

Figure 3 displays the average amount of pledge and the average number of backers by different categories. There is a large gap between the line of the average amount of pledge and the average number of backers under a category of transportation. It suggests that the average amount of pledge by a backer under a category of transportation is much higher than under the other categories. Calculation also shows that the average amount of pledge by a backer is over 399.89 US dollars under the category of transportation. On the other hand, the line of the average amount of pledge is under the line of the average number of backers under categories of Games, Politics, Comics, Video\_web and Podcasts\_blogs\_vlogs. It also suggests that the average amount of pledge by a backer under those categories is much higher than under the other categories. Additionally, the line of the average amount of pledge looks much under the line of the average number of backers under a category of Podcasts\_blogs\_vlogs, and the average amount of pledge by a backer under this category should be little. Nonetheless, since the number of projects under a category of podcasts\_blogs\_vlogs is only 26, the sample size is too small to be representative.

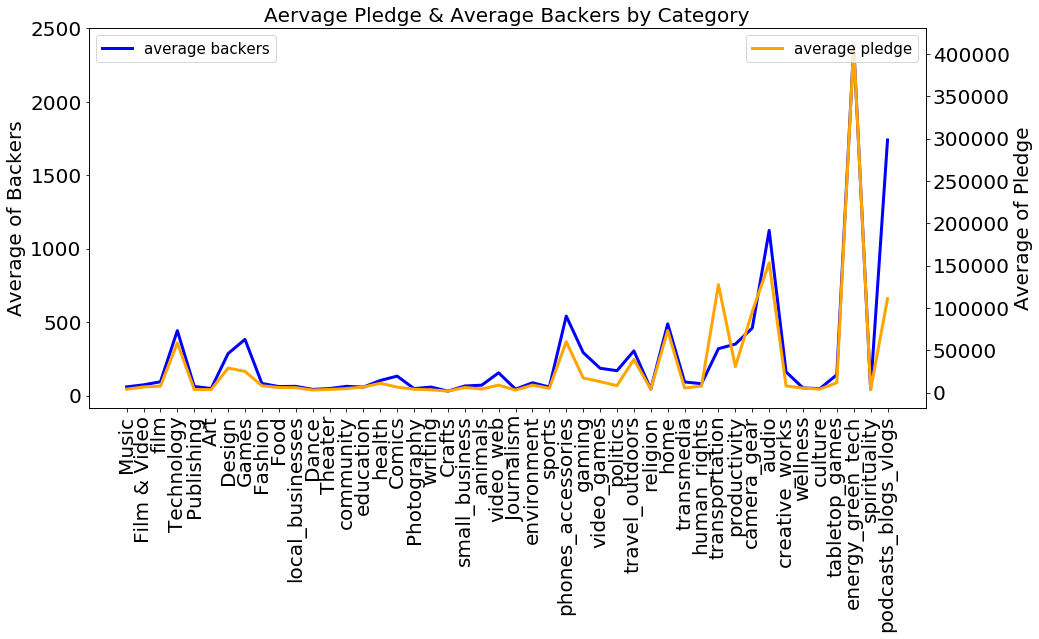


Fig 3. Average Pledge & Average Backers by Category

Figure 4, 5 and 6 respectively displays the distribution of the number of backers and the amount of pledge and goal in failed and successful projects. It can be found that the mean distribution of the numbers of backers and the amount of pledge in successful projects is higher than in failed projects. It suggests that the average number of backers and amount of pledge in successful projects are larger than in failed projects. Furthermore, the mean distribution of the amount of goal in successful projects is less than in failed projects, shows that the average amount of goal in successful projects is less than in failed projects.

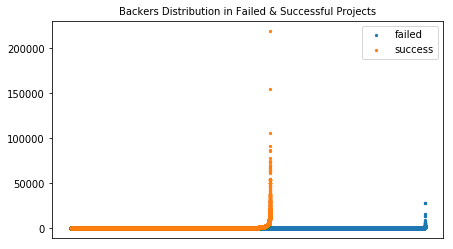


Fig 4. Backers Distribution in Failed & Successful Projects

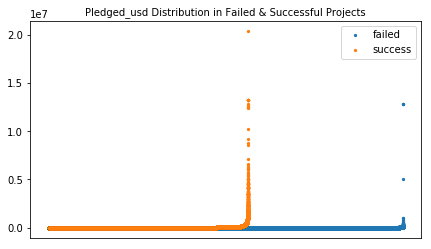


Fig 5. Pledged\_usd Distribution in Failed & Successful Projects

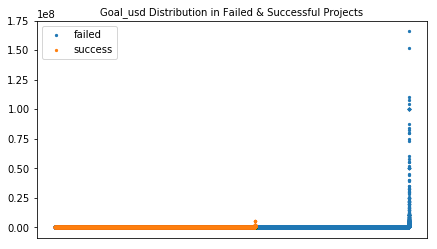


Fig 6. Goal\_usd Distribution in Failed & Successful Projects

Fig 7 presents Pearson correlation in backers, pledge\_usd and goal\_usd, which is a method to show the linear correlation between two features. The result shows that there is high correlation between pledge\_usd and backers, and the rest of features reflect low correlation in between. In further discussion, figure 8 and 9 display the correlation in features in successful and failed projects respectively. The result shows that besides the high correlation between the amount of pledge and the number of backers, the amount of goal is moderately correlated with the amount of pledge and the number of backers in successful projects. On the other hand, there is low correlation between the amount of goal and the rest of features in failed projects. Furthermore, under a category of education, which is a main category with 17025 projects, the amount of pledge is only moderately correlated with a number of backers while the amount of pledge is highly correlated with the amount of goal in successful projects. Nonetheless, there is only high correlation between the amount of pledge and the amount of goal, and the rest of features reflect low correlation in between.

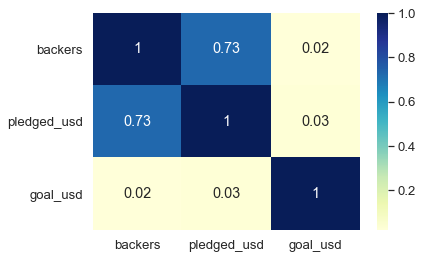


Fig 7. Pearson Correlation in Features

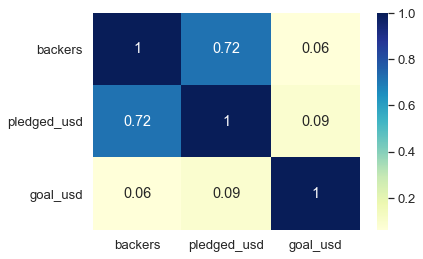
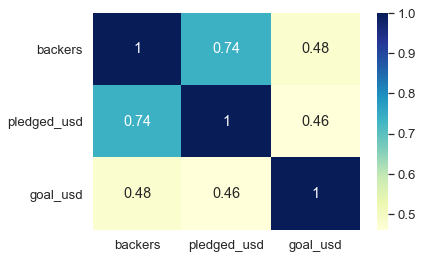


Fig 8 Pearson Correlation in Features in Fig 9 Pearson Correlation in Features in

Successful Projects Failed Projects

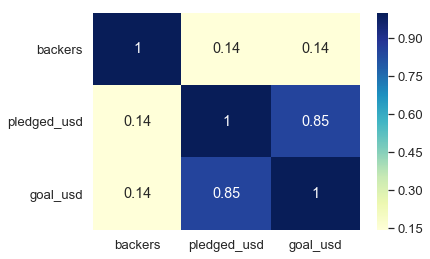
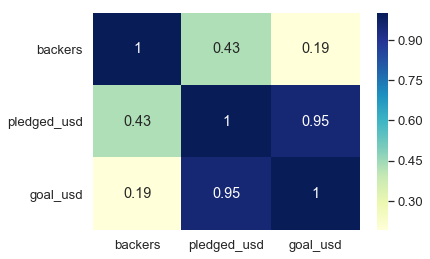


Fig 10 Pearson Correlation in Features Fig 11 Pearson Correlation in Features

in Successful Projects under the Category in Failed Projects under the Category

of Education of Education

1. Model Construction

Unfinished

1. EXPERIMENTS AND RESULTS
2. Parameters Tuning

Unfinished

1. Stacking

Unfinished

1. Comparison of Different Algorithms

Unfinished

1. Discussion

Unfinished

1. CONCLUSION AND FUTURE WORK

Unfinished

ACKNOWLEDGMENT

REFERENCE

1. Bannerman, S. (2013, 6). Crowdfunding Culture. Journal of Mobile Media, 7(01), pp. 1-30. Retrieved from Wi: Journal of Mobile Media:

<https://wi.mobilities.ca/crowdfunding-culture>

1. Belleflamme, P., Lambert, T., & Schwienbacher, A. (2010). Crowdfunding: An industrial organization perspective. In Prepared for the workshop

Digital Business Models: Understanding Strategies. Paris.

1. Canada Media Fund (CMF). (2012, 8). Crowdfunding in a Canadian Context:

<http://www.artressources.org/etatsgeneraux/CMF_SociofinancementAnglais.pdf>

1. Douglas, J., Cumming, Gael, Leboeuf, Armin, Schwienbacher, (2015), Crowdfunding Models: Keep-It-All vs. All-Or-Nothing:

<http://leeds-faculty.colorado.edu/Bhagat/CrowdfundingModels-KeppItAll-AllorNothing.pdf>

1. Freedman, D. M., & Nutting, M. R. (2015). A brief history of crowdfunding.

Retrieved from Recuperado de: <https://www.freedman-chicago.com/ec4i/History-of-Crowdfunding.pdf>

1. Gierczak, M. M.; Bretschneider, U. & Leimeister, J. M. (2014): Is all that Glitters Gold? Exploring The Effects of Perceived Risk on Backing Behavior in Reward-based Crowdfunding. In: International Conference on Information Systems (ICIS), Auckland, New Zealand.
2. icopartners.com; Kickstarter in 2015–Review in numbers. (2016). Retrieved

from icopartners.com: [http://icopartners.com/2016/02/2015-in-review/](%20http:/icopartners.com/2016/02/2015-in-review/)

1. icopartners.com; Kickstarter in 2017 – Year in review. (2018). Retrieved from icopartners.com: <http://icopartners.com/2018/01/kickstarter-2017-year-review/>
2. Kickstarter-stats. (2019). Retrieved from Kickstarter: <https://www.kickstarter.com/help/stats>
3. M.A. Arbib, Ed. (2002). The Handbook of Brain Theory and Neural Networks. Cambridge, MA: MIT Press.
4. Modwenna Rees-Mogg (2013, 9, 23). Crowd Funding: How to Raise Money and Make Money in the Crowd. Crimson Publishing
5. Sikora, Riyaz and Al-laymoun, O'la Hmoud (2014) "A Modified Stacking Ensemble Machine Learning Algorithm Using Genetic Algorithms," Journal of International Technology and Information Management: Vol. 23 : Iss. 1 , Article 1.
6. Startups.com. (2018, 5, 21). A Brief History of Crowdfunding:

<https://www.startups.com/library/expert-advice/history-of-crowdfunding>

1. The Economist. (2012, 9, 2). Putting your money where your mouse is: Crowdfunding: Artists, musicians and writers are using the Internet to aggregate lots of small donations to fund their work:

[www.economist.com/node/16909869](http://www.economist.com/node/16909869)

1. Thomas G. Dietterich (2000). Ensemble Methods in Machine Learning.
2. Tomczak, A., & Brem, A. (2013). A conceptualized investment model of

crowdfunding. Venture Capital: An International Journal of Entrepreneurial Finance, 15:4, pp. 335-359.

.