Predicting Highly Rated Crowdfunded Products

Vishal Sharma
Department of Computer Science
Utah State University, Logan, Utah, USA
vishal.sharma@usu.edu

Kyumin Lee
Computer Science Department
Worcester Polytechnic Institute, Worcester, MA, USA
kmlee@wpi.edu

Abstract—Online crowdfunding platforms have given creators new opportunities to obtain funding. Despite the popularity and success of many projects on the platforms, the quality of crowdfunded products in the market (e.g., Amazon) was not statistically and scientifically evaluated yet. To fill the gap, in this paper, we (i) compare crowdfunded products with traditional products in terms of their ratings in the largest e-commerce market, Amazon; (ii) analyze characteristics of the successful products (received \( \geq 4 \) star) and unsuccessful products (received < 4 star); and (iii) build machine learning models in three different stages, which predict whether a crowdfunded product will receive high star ratings or not. Our experimental results show that crowdfunded products, on average, received lower rating than traditional products. Our predictive models effectively identify which product will receive high star-ratings from customers on Amazon. The datasets used in this paper will be available at http://web.cs.wpi.edu/~kmlee/data.html.

I. INTRODUCTION

Crowdfunding provided creators with new opportunities to get investments from people, refine their ideas based on other users’ feedback, and have early adopters and potential brand advocates. Backers can find interesting ideas and make investment with small amount of money, and get the products early with lower price. Billions of dollars have been invested via crowdfunding platforms such as Kickstarter and Indiegogo. Kickstarter is the most popular reward-based crowdfunding platform. Some of the most successful projects on Kickstarter are Pebble Time ($20.3M pledged), Coolest Cooler ($13.3M pledged), and Kingdom Death ($12.4M pledged).

Figure 1 shows three phases of a reward-based crowdfund- ing project: (i) the fundraising phase; (ii) the reward delivery phase; and (iii) the product sale phase. In the literature, researchers focused on both fundraising and reward delivery phases. They studied whether a project will be successful in terms of raising fund [1]. They found that 9% creators failed to deliver rewards/products that they promised to their backers [2]. 35% backers did not receive rewards on time, passing estimated delivery dates [2], [3].

However, researchers did not pay attention on the product sales phase yet. According to our study, successful projects, which raised more money than their goals, did not guarantee to produce high quality products and receive high star ratings in market. For example, JamStik+\(^1\) in Figure 2, a Kickstarter project, raised $813K (16 times more than its goal), but its star rating on Amazon was low (only 2.9). It is a hard problem to predict which project will produce low quality products or which product will receive low star ratings from customers in market by using only limited online data. But, if we can predict it with a reasonably high accuracy, creators can further improve their projects, backers can support projects which will produce high quality products, and buyers can purchase high quality products as soon as it is launched in market (even without any review).

To achieve this goal, in this paper, we analyze the quality of crowdfunded and start-up products, and predict which product will receive a high or low rating based on previously available data. Since it is hard to measure quality of a product objectively, we use customers’ explicit feedbacks (e.g., star-ratings) as a way to measure quality of the product.

Recently, Amazon launched a web page called Launchpad\(^2\) where crowdfunded and start-up products are listed. By collecting and analyzing over 2 year data from Launchpad, we aim to conduct the following research objectives: RO1: compare Launchpad products with traditional products (i.e., non-Launchpad products) on Amazon in terms of customer ratings; RO2: understand characteristics of successful products (i.e., highly rated products) and unsuccessful products (i.e., lowly rated products); and RO3: build machine learning models to predict which crowdfunded product will be rated high or low when it is launched in the market.

Concretely, we made the following contributions:

- First, we compared crowdfunded products with 82 million traditional products on Amazon in terms of their rating distributions and average ratings. We found that the crowdfunded products received relatively lower rating than the traditional products.
- Second, we analyzed characteristics of successful (\( \geq 4 \) star) and unsuccessful (< 4 star) products.
- Finally, we built machine learning models in each of three stages (shown in Figure 1) to predict which crowdfunded product will be successful in market. Note in our study we use projects from crowdfunding platforms which are successfully funded.

II. DATASET DESCRIPTION

Verifying, whether a product in an e-commerce site (e.g., Amazon) was originally funded from crowdfunding platforms

\(^1\)jamstik+ The SmartGuitar: http://kck.st/2s0TLQ0
\(^2\)http://amzn.to/2ow8vpV
LAUNCH
Creators posted their projects in crowdfunding platforms.

END OF FUNDRAISING
Creators, who reached the goal, will begin making promised rewards.

MASSIVE PRODUCTION AND SALE
Creators sell their products in real marketplaces like Amazon and eBay.

Stage 1
(Fundraising phase)

Stage 2
(reward delivery phase)

Stage 3
(product sale phase)

Fig. 1. Three phases of a crowdfunding project.

TABLE I
RATING DISTRIBUTIONS OF TRADITIONAL AND LAUNCHPAD PRODUCTS.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Amazon products</th>
<th>Launchpad products</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>4,265,230 (5.2%)</td>
<td>27 (1.2%)</td>
</tr>
<tr>
<td>2.0</td>
<td>6,712,117 (8.1%)</td>
<td>108 (5.1%)</td>
</tr>
<tr>
<td>3.0</td>
<td>7,049,301 (8.5%)</td>
<td>685 (32.4%)</td>
</tr>
<tr>
<td>4.0</td>
<td>15,480,820 (18.7%)</td>
<td>961 (45.4%)</td>
</tr>
<tr>
<td>5.0</td>
<td>49,169,663 (59.5%)</td>
<td>336 (15.9%)</td>
</tr>
<tr>
<td>AVG rating</td>
<td>4.2</td>
<td>3.7</td>
</tr>
</tbody>
</table>

375 Kickstarter product dataset as Kickstarter dataset.

III. RO1: COMPARING LAUNCHPAD PRODUCTS WITH TRADITIONAL PRODUCTS

To understand gaps between the traditional products on Amazon and the products from crowdfunded websites, we analyzed the rating distributions of Amazon dataset and Launchpad dataset in macro and micro levels. Table I shows rating distributions and the average ratings in Amazon dataset and Launchpad dataset were 4.2 and 3.7, respectively.

To statistically analyze whether these datasets have different rating distributions, we performed Chi-squared test for independence. Chi-Squared test starts with defining null hypothesis (H0) and alternative hypothesis (H1):

Null Hypothesis 1: (H0) Rating is independent/not associated with a dataset.

Hypothesis 1: (H1) Rating is dependent/associated with a dataset.

Chi-squared test outputs, X-squared = 2976.4, df = 4, and p-value < 0.0001. Since Chi-squared distribution table value for df = 4 is 14.860, the result rejects our null hypothesis. It means that the rating distributions of Amazon and Launchpad datasets are not similar. The result makes sense because ratings were highly distributed between 5 and 4 star in Amazon dataset, whereas 4 and 3 star in Launchpad dataset.

To analyze average ratings in the micro level, we first chose top 5 categories according to product counts in Launchpad dataset. Then, we compared average ratings in each category of both Amazon dataset and Launchpad dataset. From Table
TABLE II
AVERAGE RATINGS IN TOP 5 CATEGORIES.

<table>
<thead>
<tr>
<th>Rating</th>
<th>[Amazon P]</th>
<th>[Launchpad P]</th>
<th>lower %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronics</td>
<td>4.01</td>
<td>3.41</td>
<td>−14.96%</td>
</tr>
<tr>
<td>Toys &amp; Games</td>
<td>4.15</td>
<td>3.97</td>
<td>−4.34%</td>
</tr>
<tr>
<td>Home &amp; Kitchen</td>
<td>4.19</td>
<td>3.76</td>
<td>−10.26%</td>
</tr>
<tr>
<td>Beauty &amp; Personal</td>
<td>4.15</td>
<td>3.77</td>
<td>−9.16%</td>
</tr>
<tr>
<td>Sports &amp; Outdoor</td>
<td>4.18</td>
<td>3.85</td>
<td>−7.89%</td>
</tr>
<tr>
<td>AVG rating</td>
<td>4.14</td>
<td>3.75</td>
<td>−9.42%</td>
</tr>
</tbody>
</table>

II, we observed that product ratings in each category of Launchpad dataset were still lower than ones in Amazon dataset by an average of −9.42%, with electronics being lowest of all by −14.96%.

Based on the analysis, we conclude that there are some gaps between traditional products and Launchpad products on Amazon in terms of quality (i.e., ratings).

IV. RO2: CHARACTERISTICS OF SUCCESSFUL AND UNSUCCESSFUL PRODUCTS

Even though the Launchpad products received lower ratings than the traditional products, 61.3% Launchpad products in Table I received a star rating ≥ 4. In this section, we are interested in analyzing characteristics of successful products (received ≥ 4 star) and unsuccessful products (received < 4 star) among the crowdfunded products. If we find distinguishing characteristics between them, we can build classifiers to predict which crowdfunded project will likely produce low quality outcome, and which crowdfunded project will likely receive a low rating from customers. Backers in crowdfunding platforms and buyers in e-commerce sites could potentially use these classifiers.

In this direction, we address two research questions: (i) Is there any positive correlation between raised money and Amazon star rating in market?; and (ii) Are there any distinguishing characteristics between successful and unsuccessful products? To answer the research questions, we use Kickstarter dataset which consists of 247 successful products (i.e., received ≥ 4 on Amazon), and 128 unsuccessful products (i.e., received < 4). Figure 3 shows the products as dots based on their pledged money and star ratings. We might assume that the larger a product’s pledged money is, the higher its star rating will be. However, it wasn’t the case in our analysis. There was no clear correlation pattern between these two properties. The Pearson Correlation between them was -0.08, showing they are not correlated. In other words, being successful in terms of raising fund in crowdfunding platforms does not mean that the creators will produce high quality products and receive high star ratings on Amazon.

To answer the second research question, we computed the mean of various properties of the successful and unsuccessful Kickstarter products on Amazon. Table III shows the list of selected properties. We observed that successful products had less number of FAQs than unsuccessful products in their Kickstarter project pages. It may indicate that backers/investors of unsuccessful products posted more concerns regarding the projects. For example, Kickstarter users asked more questions about products/projects before backing the following projects: jamStik+ and Noke3. These products received low star ratings, and were unsuccessful on Amazon. We also observed that the creators of unsuccessful products backed less number of projects than the ones with successful products, indicating that the creators of successful products are more experienced and active in the community.

In the literature, researchers found that social network plays a vital role in a project’s success [5]. We observed the same phenomena in our dataset. Creators of successful products had more Facebook friends, and were more active on Twitter. They posted more number of tweets, created more number of lists and liked others tweets. It means creators having richer and deeper social network produces high quality products.

Another interesting property is pledged money. The creators of unsuccessful products raised 59% more money than ones of successful products. However, actual products were rated low by customers on Amazon. It means raising more money

<table>
<thead>
<tr>
<th>Properties</th>
<th>Successful</th>
<th>Unsuccessful</th>
</tr>
</thead>
<tbody>
<tr>
<td>pledged money</td>
<td>$313,800</td>
<td>$28,400</td>
</tr>
<tr>
<td>[FAQs]</td>
<td>4.69</td>
<td>7.09</td>
</tr>
<tr>
<td>[comments]</td>
<td>1075</td>
<td>934</td>
</tr>
<tr>
<td>[images]</td>
<td>17.5</td>
<td>27.1</td>
</tr>
<tr>
<td>[negative comments by backers]</td>
<td>440</td>
<td>633</td>
</tr>
<tr>
<td>[projects backed by creators]</td>
<td>26.6</td>
<td>20.9</td>
</tr>
<tr>
<td>[Facebook friends]</td>
<td>773</td>
<td>359</td>
</tr>
<tr>
<td>[lists created by creators]</td>
<td>148.2</td>
<td>38</td>
</tr>
<tr>
<td>[posted tweets]</td>
<td>1,889</td>
<td>696</td>
</tr>
<tr>
<td>[tweets liked by creators]</td>
<td>1,734</td>
<td>1,397</td>
</tr>
<tr>
<td>Product Price on Amazon</td>
<td>$83</td>
<td>$107</td>
</tr>
</tbody>
</table>

3Noke: The World’s Smartest Padlock: http://kck.st/1kU8zrT
Table IV
Top 5 Features at Each Stage.

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td># of images</td>
<td># of creators</td>
<td># of creators</td>
</tr>
<tr>
<td>project description length</td>
<td># of images</td>
<td># of images</td>
</tr>
<tr>
<td>reward description readability</td>
<td># of creators’ comments</td>
<td>product price on Amazon</td>
</tr>
<tr>
<td># of backed projects</td>
<td>pledged money &amp; goal ratio</td>
<td># of Superbackers’ comments</td>
</tr>
<tr>
<td>reward description length</td>
<td># of backed Projects</td>
<td># of FAQs</td>
</tr>
</tbody>
</table>

For each stage of the project (Stage 1, Stage 2, and Stage 3), we extracted the top 5 features that were most predictive of whether the project would be successful or not. These features were primarily based on the Kickstarter page, Amazon product page, and creators' Twitter profiles. The features included the number of images, the number of creators, the reward description length, the number of backed projects, and the product price on Amazon.

Kickstarter creators' Twitter profile features: 151 out of 375 Kickstarter project creators linked their Twitter profiles. We extracted their number of tweets, number of followers, number of followees, number of favorites, and number of lists. Missing values were treated by replacing them with the mean of the respective feature.

Amazon product page features: Since we know which Kickstarter project is linked with which Amazon product, we further extracted features from an associated Amazon product page. These features consist of a category of the product, number of images, number of videos, product description length, and number of technical details. In addition, we measured a Levenshtein distance/title similarity between a product's Kickstarter title and its Amazon title. We only extracted the product page features available when it was newly created and listed to Amazon. We did not extract any feature from comments and reviews associated with the Amazon product page because we assume that our predictive model (which will be described shortly) will predict whether an Amazon product will be successful or not once its product page is just created.

V. RO3: BUILDING PREDICTIVE MODELS

A. Feature Engineering

First of all, we describe our proposed features which will be used to build predictive models. The features are grouped by four categories: (i) Kickstarter project page features; (ii) Kickstarter creators' profile features; (iii) Kickstarter creators' Twitter profile features; and (iv) Amazon product page features.

Kickstarter project page features: These features were extracted from each product's associated Kickstarter page. They consist of a project goal, pledged money, number of images, number of videos, number of FAQs, number of comments, number of rewards by creators, number of backers in least rewards, number of backers in maximum rewards, project description length, reward description length, a percentage of negative comments associated with the project, Coleman Liau readability scores [8] of the project page and reward descriptions, a ratio of pledged money to the goal of the project, and number of comments from Superbackers4 The percentage of negative comments was calculated by a sentiment analyzer [7].

Kickstarter creators' profile features: These features were extracted from creators' profile. They consist of number of backed projects, number of created projects, number of linked external websites, number of creators (e.g., a project may be created by multiple people), is the account verified?, is Facebook connected?, and number of Facebook friends.

### Footnote

4Superbackers are users who have supported more than 25 projects with pledges of at least $10 in the past year.
to reviews because the reviews are not available yet when the Amazon product page is just launched.

In this experiment, we used Kickstarter dataset. To make sure all the features have distinguishing power between successful and unsuccessful products, we conducted feature selection by measuring mean decrease impurity from Random Forest. Table IV shows top 5 features at each stage. In particular, in the stage 3, five most important features were # of creators, # of images on Amazon, # of technical details, Similarity between a product’s Kickstarter title and its Amazon’s title based on Levenshtein distance to reviews because the reviews are not available yet when the Amazon product page is just launched.

In this experiment, we used Kickstarter dataset. To make sure all the features have distinguishing power between successful and unsuccessful products, we conducted feature selection by measuring mean decrease impurity from Random Forest. Table IV shows top 5 features at each stage. In particular, in the stage 3, five most important features were # of creators, # of images on Amazon, # of technical details, Similarity between a product’s Kickstarter title and its Amazon’s title based on Levenshtein distance.

Next, we also analyzed the Random Forest model’s partial dependence plots [9] which reveals which feature positively or negatively affected the model. Especially, we focused on finding negatively affecting features via the plots so that we can further improve our model. By running partial dependence plots, we found that four features – number of rewards, pledged money, number of followers, number of created projects – negatively affected the model. After removing these features, our model achieved 0.761 accuracy. Figure 4 shows two-way partial dependence plots of number of created projects and number of rewards. Pairs of the features negatively affected the model.

**VI. RELATED WORK**

In this section, we summarize some of the prior work related to rating prediction and crowdfunding. Researchers have studied Amazon reviews for rating prediction which uses text of reviews, and/or annotations of the reviews. For example, Tang et al. [10] developed a neural network-based method that uses both reviews and author information. Gupta et al. [11] predicted rating by using supervised learning. However, our method does not require reviews. Instead, it leverages information from a crowdfunding website, and predicts whether a product will be successful or not as soon as the product page is created on Amazon. This makes our work novel and first of its kind.

Crowdfunding platforms have been studied widely in recent years [12], [13]. There have also been several works related to the prediction of project success. Researchers built classifiers...
based on directly extracted features from Kickstarter to predict the success of a project (i.e., achieving equal to or greater than its goal) [14]. Etter et al. [15] proposed a method to predict project success by using direct information and social features. Mitra et al. [16] studied a corpus of 45k projects and proposed novel text features for the prediction.

Researchers analyzed several factors which affects project success. Joenssen et al. [17] showed that the timing and communication were important factors affecting project success. Xu et al. [1] extensively studied how project updates are highly correlated to project success. Tran et al. [18] showed different factors associated with making a project successful. Geographical factors, project updates and rewards have also helped in understanding projects success on crowdfunding websites [1, [19], [20]. Other researchers [5], [21], [22] analyzed the social media’s and social communities’ role in raising fund. Mollick et al. [23] studied dynamics of success and failure of crowdfunding products, and found that social/personal networks and project quality were associated with the crowdfunding success.

Recommendation of projects to backers, and backers to creators have also been studied widely. In [24], authors built a recommender system to recommend potential backers to creators. In [22], authors build another recommender system to recommend potential projects to backers. Rakesh et al. [25] proposed a recommendation model to recommend projects to group of investors. However, researchers did not pay attention to the performance of crowdfunded products in real markets. To complement the prior work, we analyzed characteristics of successful (i.e., rated high stars) and unsuccessful (i.e., rated low stars) products, and then built machine learning models to predict success of a product. To our knowledge, we are the first one to explore this topic and area of research.

VII. Conclusion

We found that Launchpad products, on average, received lower ratings than traditional products on Amazon. We also found that there were distinguishing properties between successful and unsuccessful products. Based on the analysis and observation, we built predictive models which predict whether a product will receive high ratings or not when it is launched in the market. Random Forest based classifier outperformed the other models, and improved accuracy by 15% compared with the baseline. In our continuing research, we are interested to expand our Launchpad dataset up to 10,000s, add new features such as clustering information and a product’s difficulty level, and consider other explicit feedbacks (e.g., # of reviews and Amazon sales rank) as ways to measure qualify of products. We are also interested to expand our work to other crowdfunding platforms and e-commerce websites.

Acknowledgment

This work was supported in part by NSF grant CNS-1755536, Google Faculty Research Award, Microsoft Azure Research Award, and Nvidia GPU grant. Any opinions, findings and conclusions or recommendations expressed in this material are the author(s) and do not necessarily reflect those of the sponsors.

References