Plead or Pitch? Predicting Performance of Kickstarter Campaigns

Nihit Desai, Raghav Gupta and Karen Truong
Department of Computer Science, Stanford University

ABSTRACT

Kickstarter is a crowdfunding portal where entrepreneurs and artists look for capital for their projects. In contrast to prior work, which relies largely on features from the project's intermediate/final performance to predict success/failure, we tackle the potentially more useful task of predicting success based only on the initial description of the project - something that the project creator has full control over.

We focus mainly on the language of the project description- in particular, phrases which are predictive of success- and their psycholinguistic qualities - in addition to other metadata. We report performance gains over existing approaches when using only the initial project description

DATASET

After filtering, our dataset consists of approximately 26K Kickstarter projects, collected over a period of 8 months, with daily snapshots of each project's page over the course of the campaign to account for modifications made to the campaign by the project creator.

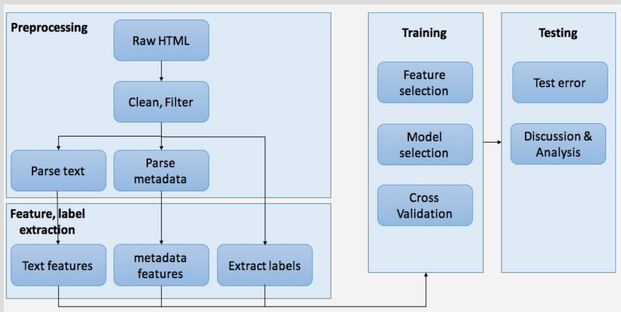
Kickstarter works on an "all-or-nothing" funding basis, where a project receives pledged money only if it meets/exceeds its funding goal by the campaign's end date. We hence consider "funded" and "not funded" as our labels for classification.

Dataset Summary:

Total no. of campaigns	26543
No. of successful campaigns	7862

Metric ↓ Outcome →	Successful	Failed
Campaign duration	27.1	21.2
Goal amount	$5747	$19344
# Sentences	24.5	23.6
# Rewards offered	8.73	7.52

STEPS



The diagram above illustrates our workflow for this project.

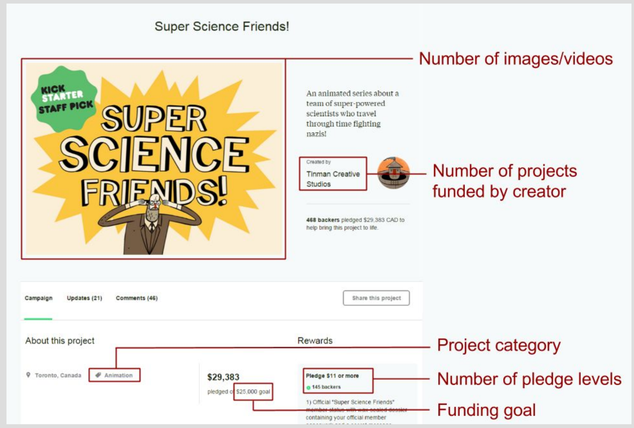
REFERENCES

- Etter, Vincent, Matthias Grossglauser, and Patrick Thiran. "Launch hard or go home!: predicting the success of kickstarter campaigns." *Proceedings of the first ACM conference on Online social networks*. ACM, 2013.
- Pennebaker, James W., Martha E. Francis, and Roger J. Booth. "Linguistic inquiry and word count: LIWC 2001." *Mahway: Lawrence Erlbaum Associates* 71 (2001): 2001.

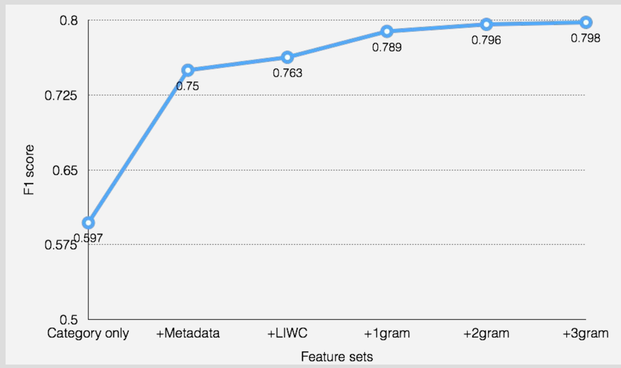
FEATURES

Our feature set has two major components -

1. **Linguistic features**, which include
   a. Uni, bi and tri-grams from the project description and the section on risks, for predictive phrases. We selected the top 600 n-grams using SVD
   b. Psycholinguistic features (LIWC - Linguistic Inquiry and Word Count) for categories (cognitive, inhibition)
   c. Sentiment scores from comments on project page, using Stanford CoreNLP
2. **Non-text metadata**, such as (but not limited to)



Starting with a simple feature set that predicts success based only on the project category, we **progressively add features** and track performance for a given model (F1 score averaged over 5-fold CV, L2 logistic regression, balanced class weights):



RESULTS

- We use **training** and **dev sets** for experimenting with different estimators and model selection
- We **adjust class weights** (since dataset unbalanced): a "most common label" baseline will have accuracy of ~0.5
- For these experiments, we use project snapshots as on **first day of the campaign**
- SVM (RBF kernel) and Logistic Regression (L2 regularization) perform best, SVM being slightly better.

	Successful		Failed		Overall		
	P	R	P	R	P	R	F1
Multinomial NB	0.65	0.58	0.8	0.84	0.75	0.76	0.75
SGD	0.71	0.48	0.68	0.85	0.70	0.69	0.68
Decision Tree	0.54	0.55	0.84	0.82	0.77	0.75	0.76
Ridge Classify	0.72	0.65	0.82	0.87	0.78	0.79	0.78
Logistic Reg.	**0.74**	0.61	0.80	**0.88**	0.78	0.78	0.78
$SVM_{polynomial}$	0.53	0.59	0.84	0.81	0.76	0.75	0.75
SVM_{RBF}	0.71	**0.65**	**0.84**	0.87	**0.79**	**0.80**	**0.79**

- We use SVM with RBF kernel to see how well can we predict success on **test set** as the campaign progresses.
- Performance on first and last day of the campaign give lower and upper bounds:

	Successful			Failed			Overall		
	P	R	F1	P	R	F1	P	R	F1
First day	0.72	0.63	0.66	0.84	0.87	0.84	0.79	0.79	**0.79**
Last day	0.88	0.72	0.79	0.86	0.94	0.90	0.86	0.86	**0.86**

- At last, we qualitatively analyze n-gram features to see the types of phrases most predictive of success on Kickstarter

Social	friends, community, friendship
Transact	free shipping, in return for, you receive, early bird
Emotion	dream, passion, impact, believe that, inspired to start

- Turns out there is a bit of both pitching & pleading in successful projects

Pitch	why support, conversation, practical, if you decide
Plead	help, grateful, support, funds will cover