



Solving Kickstarter

Predicting Campaign Success with Support Vector Machines

Some Background...

What is Kickstarter?

- ❑ A **crowdfunding platform**
- ❑ An alternative, public method of **gathering investments for an idea**
- ❑ Founder Perry Chen: “A middle ground between patronage and commerce”

More Background...

How does Kickstarter work?

- ❑ Projects are created, along with a **deadline and a minimum funding goal**
- ❑ If a **goal is not met**, no funds are collected, and the **project is deemed a failure**
- ❑ Ideally, successful project creators use the investments they receive to deliver the promised product

Our Data

Our algorithm made use of a list of over 300,000 recent kickstarter campaigns and their various attributes in a tabular format

Our Goal

Predicting Kickstarter Campaign Success

- ❑ Utilize campaign characteristics to observe what makes a campaign successful
- ❑ Find patterns, consistencies etc. in these attributes in order to accurately predict campaign success with applicable attributes
- ❑ Create a tool that can effectively predict outcome of any Kickstarter campaign

The Problem

Not all features or attributes are available before campaigns are decided

Our Approach

Titles - An Unconsidered Field

- ❑ Projects considering Kickstarter categories have been carried out before - but the titles of the campaigns have been largely ignored
- ❑ The **positivity score** of a title can be calculated and utilized
- ❑ The **reading ease score** of a title can be calculated and utilized
- ❑ Do the positivity score and the reading ease score of a campaign significantly impact success?

Implementation

Tailor the dataset to fit our needs
and train a model to derive
accuracy

Tailoring our Dataset

Converting Textual Components

- ❑ Converted all **textual components to a numeric value or to dummy values** to feed into algorithm
- ❑ **Sentiment Analysis Score** - calculated the net positivity score of the campaign title
- ❑ **Flesch Kinaid Score** - calculated the net reading ease of the campaign title

Our Implementation

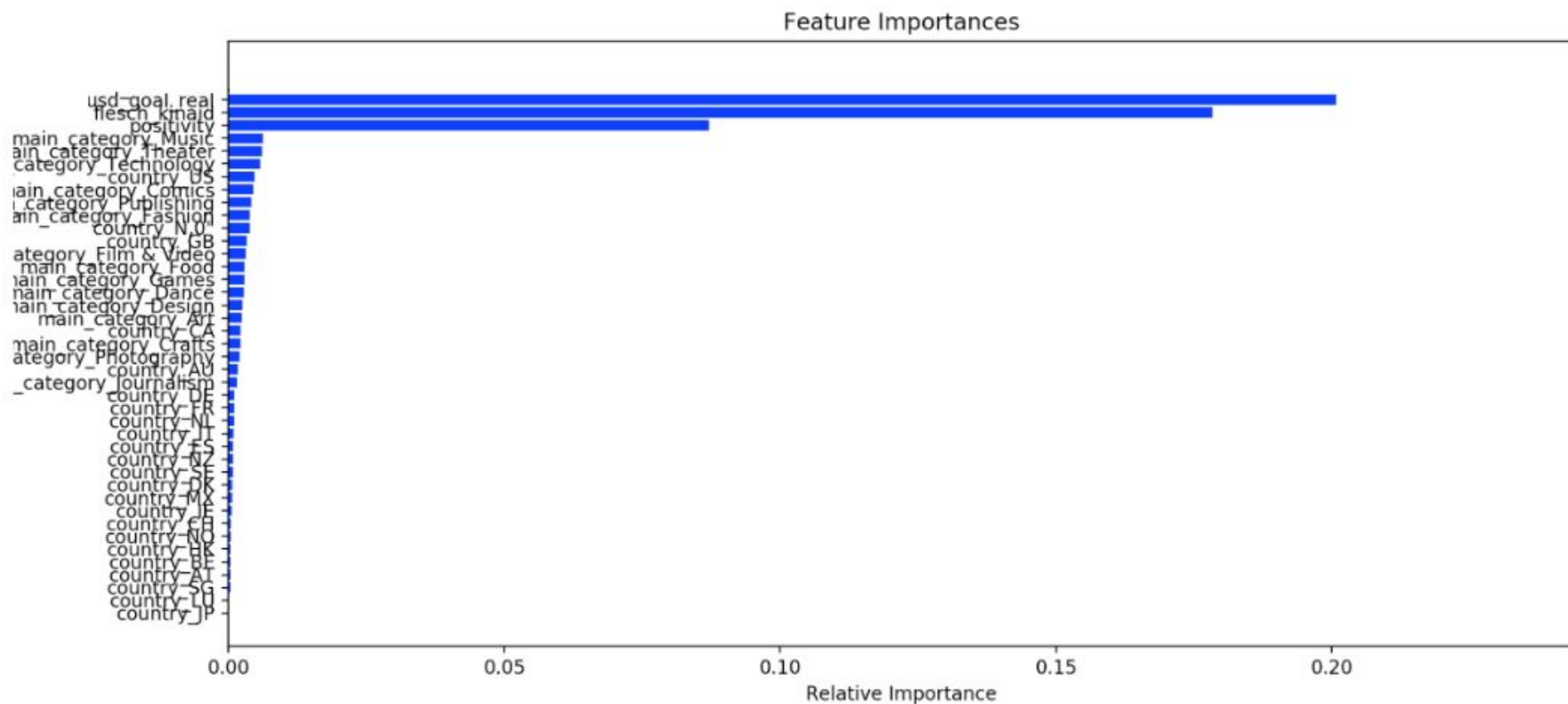
Machine Learning Models

- ❑ **Linear Support Vector Classification** - it has flexibility in the choice of penalties and loss functions and should scale better to large numbers of samples
- ❑ **Random Forest Classifier** - fits a number of decision tree classifiers on various sub-samples of the dataset to calculate **Feature Importance**

Results

Reading Ease and Positivity of a campaign's title play a role in predicting campaign success

Feature Importance



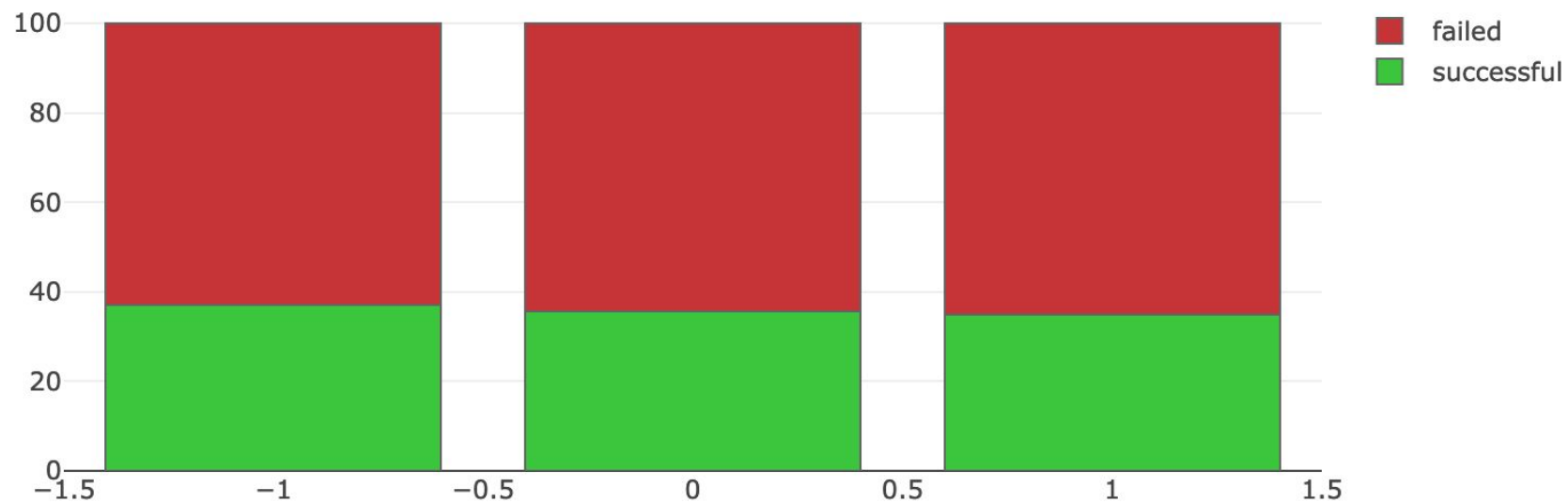
Model Accuracy

66%

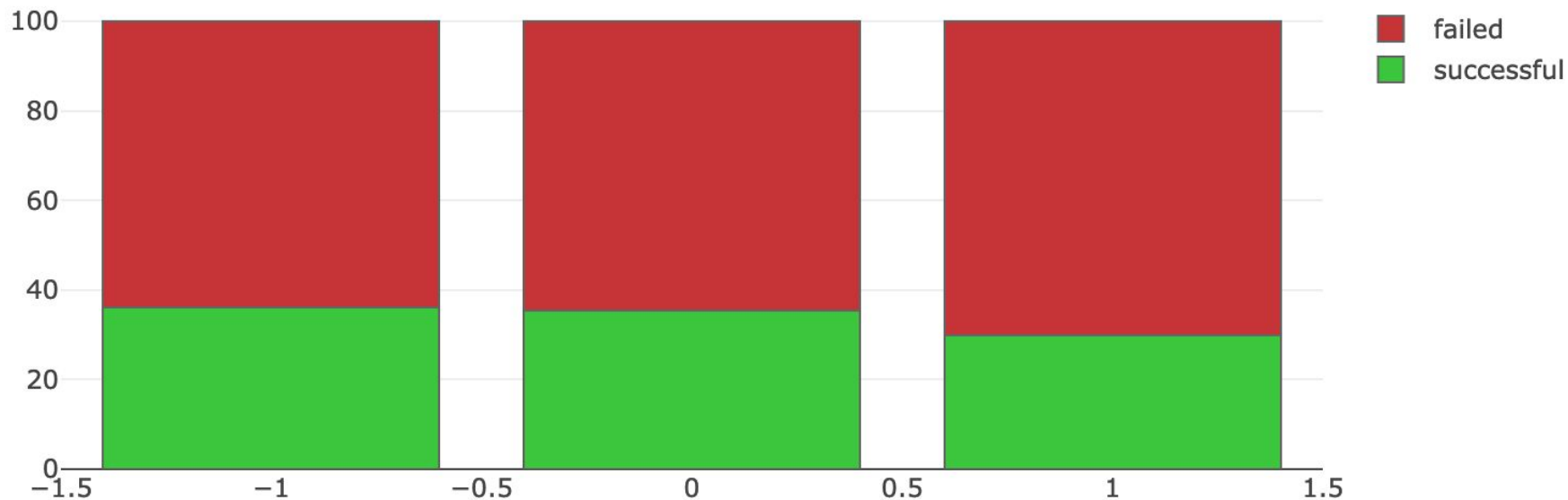
Accuracy in predicting campaign success or failure

In Conclusion

% of successful and failed projects by positivity



Flesch Kinaid Score vs. Success



Thank you!

Any Questions??
