

PROJECT NO. 2

MACHINE LEARNING

Data: Kickstarter

neuefische
Data Science Bootcamp 2021
Cologne

Christian **Dischke**
Burak **Özbagci**
Matthias **Schmidt**
2021-06-29

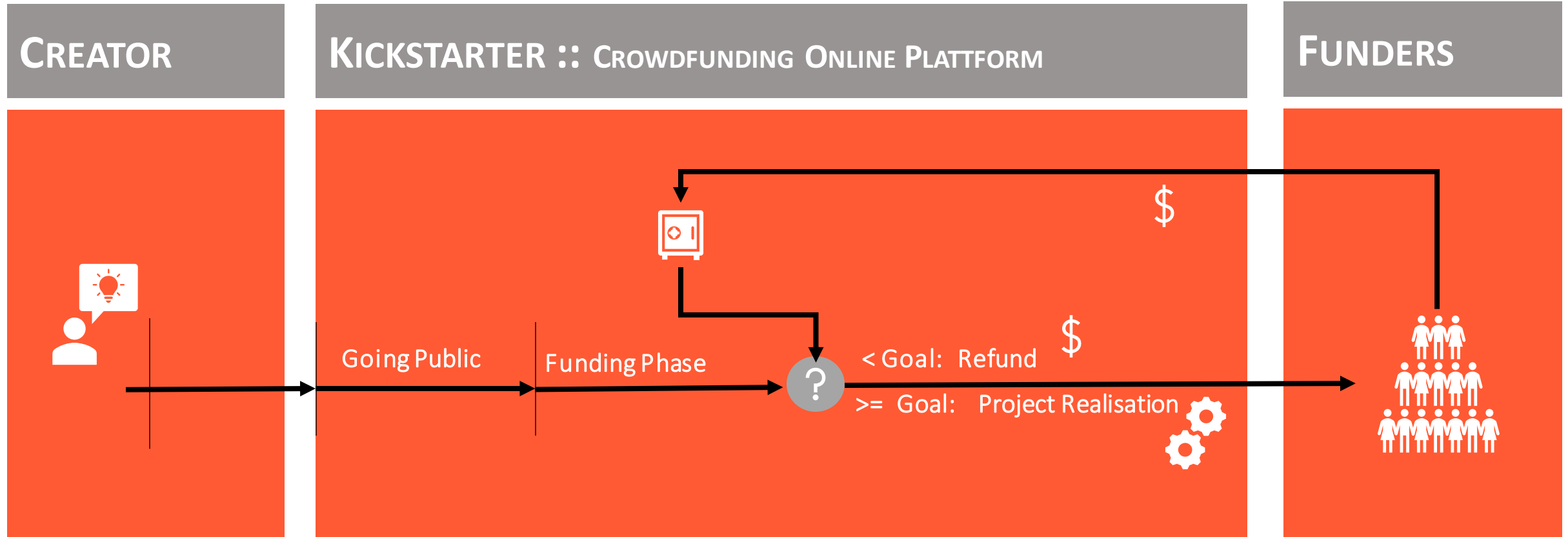
01

Introduction



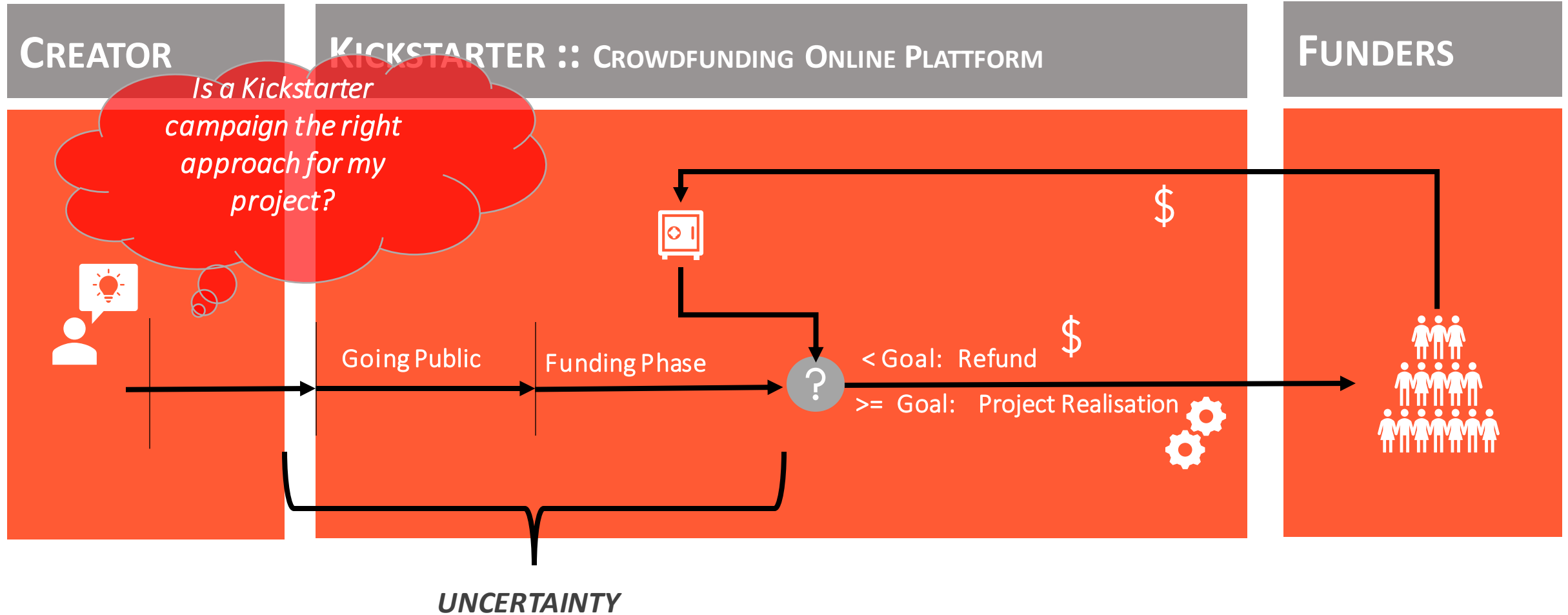
KICKSTARTER :: ALL OR NOTHING

SIMPLIFIED PROCESS



KICKSTARTER :: ALL OR NOTHING

SIMPLIFIED PROCESS



OUR TASK

U

“

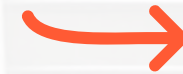


Discovery and Strategy

We take the time to create a solid foundation for your campaign through our discovery and strategy sessions so we can fully understand your product and how best to market your project. We also offer videography and photography consultations and services if needed.

”

Ref. Kickstarter Marketing Services for Crowdfunding



Can we predict the success of a Kickstarter campaign before its creation?






02



Approach

What we are allowed to know

PREDICTION BY CREATORS PERSPECTIVE

CONSTRAINTS AND CONTROLLABLES

Dataset 	
Constraints	Controllables
<ul style="list-style-type: none">• Product Category• Creators• Creators Location<ul style="list-style-type: none">• Nation• State• City	<ul style="list-style-type: none">• Campaign Goal• Campaign Dates<ul style="list-style-type: none">• Creation• Launch• Deadline• Campaign description

Domain Knowledge 
Context
<ul style="list-style-type: none">• Location aggregation by aspects of:<ul style="list-style-type: none">- Economic area- Geographical proximity- Cultural proximity• Campaign<ul style="list-style-type: none">• Durations• Phase ratios



03

Prediction Results of 3 Machine Learning Models

MACHINE LEARNING RESULTS



What is the **overall** performance of our models?

Machine Learning Model	Time [s]	Accuracy
Model 1	150	81% ✓
Model 2	3	80%
Model 3	< 1 ✓	66%
Benchmark	-	56%



How accurate are **successful** campaigns predicted?

Machine Learning Model	Precision
Model 1	86%
Model 2	85%
Model 3	99% ✓

04

Summary



SUMMARY OF RESULTS

AdaBoost

1

- Highest **overall** performance, but quite slow.

Logistic
Regression

2

- Lower performance as AdaBoost
- Fast. More applicable for real-time application

Naive Bayes

3

- Quite vague overall performance



It's a tradeoff between time and overall performance.

*Nevertheless, **LR** is significantly fast*



We won't recommend Naïve Bayes.



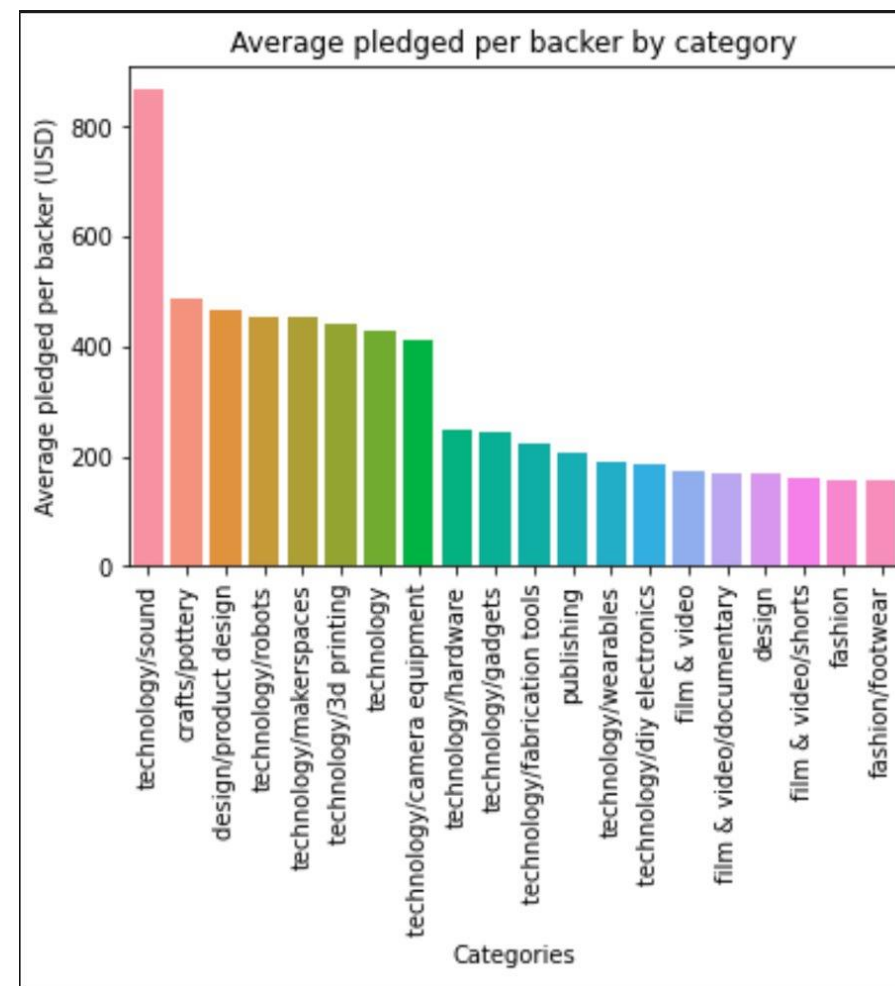
05

Critical Appraisal



CRITICAL APPRAISAL

- Up-to-date
 - Newest entry: 14 March 2019
- Biggest contributors for success: Hype and fresh idea
- Limitations of the results
 - No relevancy for content of project





06

Outlook

OUTLOOK

Data Set

Review of Constraints & Controllables

Deeper data exploration to generate action recommendations for creators.

Additional Data Mining

Data about the Creator

Scope of additional experience:
Funder and Co-creator.



✓ Thomas Franken

🔒 Letzte Anmeldung 22. Jun 2021

📘 Mit Facebook verbunden

🎨 Erstes Projekt • 7 unterstützt

Mitarbeiter bei diesem Projekt

👤 BackerKit Collaborator

Kontaktieren

Domain Knowledge

Our Hypothesis

Consider US Pay Cycles: 15 and 30 every month





07

Discussion