



Predicting outcomes in crowdfunding campaigns with textual, visual, and linguistic signals

Jermain C. Kaminski · Christian Hopp

Accepted: 13 March 2019 / Published online: 9 July 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract This paper introduces a neural network and natural language processing approach to predict the outcome of crowdfunding startup pitches using text, speech, and video metadata in 20,188 crowdfunding campaigns. Our study emphasizes the need to understand crowdfunding from an investor’s perspective. Linguistic styles in crowdfunding campaigns that aim to trigger excitement or are aimed at inclusiveness are better predictors of campaign success than firm-level determinants. At the contrary, higher uncertainty perceptions about the state of product development may substantially reduce evaluations of new products and reduce purchasing intentions among potential funders. Our findings emphasize that positive psychological language is salient in environments where objective information is scarce and where investment preferences are taste based. Employing enthusiastic language or showing the product in action may capture an individual’s attention. Using all technology and design-related crowdfunding campaigns launched on Kickstarter, our study underscores the need to align potential consumers’ expectations with the visualization and presentation of the crowdfunding campaign.

Keywords Startups · Crowdfunding · Pitch · Machine learning · Neural network · Natural language processing

JEL Classification L26 · G2 · C45 · C55

1 Introduction

“Our inviolable uniqueness lies in our poetic ability to say unique and obscure things, not in our ability to say obvious things to ourselves”—(Rorty 1979, 123)

Over the past years, crowdfunding is increasingly chosen as a gateway to overcome the financial bottleneck for early-stage ventures and new venture development processes. In crowdfunding, many small investors can contribute to a proposed new product before the product hits the market. Contributions can range from a few dollars to substantial investments into high-technology tools. The financial vehicle has been exceptionally well perceived in areas such as 3D printing,

J. C. Kaminski (✉)
School of Business and Economics, Maastricht University,
6211 Maastricht, LM, Netherlands
e-mail: j.kaminski@maastrichtuniversity.nl

C. Hopp
RWTH Aachen University, TIME Research Area, 52182 Aachen,
Germany
e-mail: hopp@time.rwth-aachen.de

virtual reality, do-it-yourself electronics, or wearables, and may also foreshadow more general demands in these industries (Mollick 2014b; Allison et al. 2015; Ahlers et al. 2015; Kaminski et al. 2017). Despite crowdfunding backers exhibiting expert-like expertise in technological areas, backers are plagued by uncertainty surrounding campaign feasibility and crowdfunders' technical expertise. At the inception of a crowdfunding campaign, many ventures have at best completed slightly more than half of the proposed milestones in new product development (Stanko and Henard 2017). Crowdfunding therefore presents unique challenges, as product possession is temporally distant there is a long gap between product possession and the time of the amount contributed to the campaign (Mollick and Kuppuswamy 2014a). Mollick and Kuppuswamy (2014a) find that more than 75% of successfully funded Kickstarter projects deliver products later than expected (i.e., only 23–25% are on time). The study also finds that project size and increased expectations around highly popular projects are related to delays. Larger projects suffer much longer delays than smaller projects, especially in the case of over-funded campaigns.

Consequently, potential backers in crowdfunding are looking for potential cues to reduce uncertainty and predict new venture success when making their capital contributions (Mollick 2013; Ahlers et al. 2015). One way for innovators to overcome this uncertainty is to signal competence trust, arising from expectations about the competence of the innovator, to create a higher receptivity among potential contributors (Sako 1992). Prior work has shown that impression management (Parhankangas and Ehrlich 2014), competence signaling (Gafni et al. 2019), and persuasion (Allison et al. 2017) may all affect crowdfunding positively. However, prior work has found mixed evidence on the role of visual and textual cues. While Parhankangas and Renko (2017) find that commercial entrepreneurs need to primarily focus on product, or firm and entrepreneur-related signals in their textual descriptions, other work shows that in low attention states visual cues work best, while textual information become only relevant if a high attention has been triggered previously (Allison et al. 2017). Hence, the effectiveness of a crowdfunding campaign pitch is inextricably linked to the various media involved.

In order to increase their funding success, we believe that project owners have a propensity to strategically use project descriptions and video pitches as marketing tool to influence potential backers' contribution decisions. In

this respect, campaign information in crowdfunding in general and video content in particular can be considered as comprehensive signals. The information available to potential backers can assist to form expectations and may induce the belief that the campaign founder possesses the relevant skills and knowledge to perform the project task for a shared mutual benefit. Information shared tacitly through videos and campaign information can, therefore, reduce the perceived performance risk of crowdfunding campaigns and should lead to higher capital endorsements. Information helps to overcome "the shadow of the future" (Axelrod and Hamilton 1981) and to reduce information asymmetries (Akerlof 1970).

Unfortunately, prior research has failed to comprehensively address the interplay of various forms of signals and cues available in crowdfunding. Most research is still embedded in survey-driven or experimental data and has not taken advantage of newer methods to encompassingly tackle the challenges that the large repository of crowdfunding data represents. At the same time, much progress has been made toward artificial intelligence, using machine learning systems that are trained to replicate the decisions of human experts (LeCun et al. 2015). These expert systems (Hayes-Roth et al. 1983) tackled challenging domains in terms of human intellect, such as image recognition (He et al. 2016), language translation (Wu et al. 2016), medical image classification (Esteva et al. 2017), mastering board games Go, Shogi, or Chess (Silver et al. 2016, 2017, 2018), playing computer games (Mnih et al. 2015), and achieved or exceeded human-level performance (LeCun et al. 2015). A comparison of the annual publishing rates of different categories of academic papers, relative to their publishing rates in 1996, shows that the number of papers on artificial intelligence increased more than nine-fold (Shoham et al. 2017, 10). Likewise, in the economics domain, machine learning techniques and methods on causal inferences entered the econometric toolbox (Varian 2014; Athey and Imbens 2017; Kleinberg et al. 2017; Mullainathan and Spiess 2017; Belloni et al. 2014).

In the following, we, therefore, explore computational techniques to predict crowdfunding campaign success based on the informational cues provided within campaign text, speech, and videos. Advances in data processing and machine learning allow new ways of analyzing data and may have profound implications for empirical testing of lightly studied, yet complex, empirical relationships. That being said, we propose the idea

that new forms of internet-mediated capital, such as crowdfunding, provide comprehensive and potentially computable signals to predict outcomes or provide recommendations. For instance, crowdfunding could be considered as perhaps the biggest open laboratory to study the interaction of inventors and investors at large scale.

In this research, we propose a novel method that combines neural networks and text-mining to identify features of successful crowdfunding projects, using transformed text, speech, and video content. Using text, speech, and video object-related meta-data in 20,188 crowdfunding campaigns, our analysis employs natural language processing techniques and neural network models to predict the success of crowdfunding campaigns. Based on word and paragraph vector models of text, speech, and video information, a feature-union model achieves a prediction accuracy of 73% in explaining campaign success or failure. Besides, we derive dialectic particularities in text, speech, and video characteristics that determine whether campaigns are more likely to be successful. Our study emphasizes the need to understand crowdfunding from a consumer's and future investor's perspective. Linguistic styles in crowdfunding campaigns that aim to trigger excitement, or are aimed at inclusiveness, are better predictors of campaign success than firm-level determinants. At the contrary, higher uncertainty perceptions may substantially reduce evaluations of new products and reduce purchasing intentions among potential funders. Our findings emphasize that positive psychological language is salient in environments where objective information is scarce and where investment preferences are taste based. We believe that our work helps to challenge and to reconsider prevailing theoretical assumptions about the prediction of entrepreneurial outcomes.

2 Methodology

We follow along the line of prior research that pays attention to the textual and linguistic context of crowdfunding campaigns. Early work here focused on the prediction of campaign success using text-mining features from project descriptions (Greenberg et al. 2013). Researchers used decision tree (DT) algorithms and support vector machines (SVC) to train a machine learning classifiers on explaining campaign success (Greenberg et al.

2013). Models achieved 68% accuracy with their respective datasets, an improvement of roughly 14% over the related baseline. More recent research focuses on the predictive power of project description content, specifically the words and phrases project creators use (Mitra and Gilbert 2014). In here, linguistic features extracted from project descriptions were combined with other campaign features to predict crowdfunding success. Tools such as Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al. 2001; Tausczik and Pennebaker 2010) infer psychologically meaningful styles and social behaviors from unstructured text (Mitra and Gilbert 2014; Desai et al. 2015; Kaminski et al. 2017; Parhankangas and Renko 2017). Mitra and Gilbert (2014) conclude that the language used in the project has a surprisingly high predictive power, accounting for about 59% of the variance around successful funding. More recent considerations of n-gram features in language employ time-variant models, i.e., data related to the beginning and end of campaigns for the prediction, showing an increased accuracy of predictions, with more available information toward the end of a campaign (Desai et al. 2015). Similar research investigated the n-gram features of "lead users" (von Hippel 1986) on crowdfunding platforms (Kaminski et al. 2017).

Based on the theory of the Elaboration Likelihood Model (ELM) (Petty and Cacioppo 1986; Bhattacharjee and Sanford 2006), Du et al. (2015) study the influence of project descriptions on crowdfunding success. Using constructs such as argument quality (number of words, readability regarding the Gunning Fog Index, sentiment ratio) and source credibility (previous campaign track record), the model predicts funding success with an accuracy rate of about 71–73%. Using campaign description text data only, Lee et al. (2018) present work building upon sequence-to-sequence (seq2seq) deep neural network models with an average 76% prediction accuracy on the first day of project launch.

Lastly, other approaches focused on contextual variables such as the social network activity of campaigns to predict funding success. For instance, the size of the social network of founders positively influences project success (Mollick 2014b). Social media activity explains some 75% variations in campaign success when conditioning on early project stages (Lu et al. 2014).

More recently, studies already began to employ machine learning classifiers to predict the temporal backing patterns using project-based information and social features, obtained from Twitter (Etter et al. 2013; Li et al. 2016; Tran et al. 2016) and backer-network graphs (Etter et al. 2013). Using a k-nearest neighbors (kNN) classifier and a Markov Chain, Etter et al. (2013) predicted the trajectories of money pledged to campaigns. Drawing upon a dataset of 16,042 campaigns, a logistic regression and linear SVC estimator reach an accuracy of more than 76% (a relative improvement of 4%), 4 hours after the launch of a campaign (Etter et al. 2013). A similar stream of research focuses on the social dimension of campaigns considers the sentiment from user comments on campaign updates (Desai et al. 2015; Lai et al. 2017). Results suggest that the text sentiment and quality in comments one week after launch are very predictive for a campaign's outcome.

As Greenberg et al. (2013), Hui et al. (2013), and Yuan et al. (2016) conclude, prediction models can be used to give feedback on proposed campaigns or as a tool to match projects with potential investors (An et al. 2014).

Notwithstanding these contributions, there is a dearth of studies considering actual *speech* content and visual campaign narratives to predict crowdfunding success. For instance, analyzing the linguistic style of crowdfunding pitches enables to conclude about revealed emotions and speech characteristics of creators (Kim et al. 2016), to distinguish between social or commercial entrepreneurs (Parhankangas and Renko 2017), or to separate conventional from "lead users" (von Hippel 1986) induced crowdfunding campaigns (Kaminski et al. 2017; Oo et al. 2018).

Concerning the analysis of *video* content, only standard approaches have hitherto been used for the evaluation of qualitative content and to measure the subjective perception of crowdfunding videos. Analyses mainly relate to the storyline and social construction (Doyle et al. 2017), perceived innovativeness, passion, preparedness, video quality, product appeal, perceived effort (Koch and Cheng 2016; Chan and Parhankangas 2017; Dey et al. 2017), and lead user appearance (Kaminski et al. 2017; Oo et al. 2018).

Our work, therefore, differs from the previous studies in four important ways:

1. First, we consider combined text, speech, and video information in our analysis. We, therefore, believe the approach covers the full spectrum of human-like campaign experience, including the processing of text, speech, and visual appeal.
2. Second, we employ proven machine learning methods to predict crowdfunding success. Our work considers Doc2Vec (see Section 3.5) paragraph vectors to model the extent to which language predicts campaign success.
3. Third, we focus on a homogeneous product category sample and restrict our analysis to technology-related products in the Kickstarter categories *Technology* and *Product Design* only. In doing so, we are strongly convinced that these two categories and inherent product presentations mostly signal "startup character." Many technology product campaigns approximate more well-known startup companies, as they signal the goal of becoming long-lasting projects, i.e., corporations that emerge with the support of the crowd (Mollick 2014b; Cordova et al. 2015; Parhankangas and Renko 2017). Indeed, research by Mollick and Kuppuswamy (2014a) on reward-based crowdfunding indicates that more than 90% of successful projects remained ongoing ventures and that 32% of all these reported yearly revenues of over \$100,000 a year since the Kickstarter campaign. Mollick (2015) further finds that only about 9% of all projects fail to deliver. Hence, our findings are potentially generalizable to the broader set of de novo firms that are founded and carry implications for the marketing and promotion of these ventures alike.
4. Fourth, we marginally improve the prediction accuracy by including information in speech content and video content. A combined approach of text, speech, and video content will, therefore, shield against a loss in information and provide more accurate estimates, as for some campaigns, descriptions are entirely encoded in images (Desai et al. 2015).¹ Similarly, we can account for all types of information processing preferences that potential backers may have. Let that be learning through reading or by indulging in video-related content. Altogether, our model covers a variety of text,

¹ Many Kickstarter campaigns insert images in their campaign pitch instead of raw HTML text. While creators use images to embed information about the team, details of the product, prototype development, or stretch goals, information is inaccessible for a comprehensive analysis and will likely yield biased results, if not taken into account explicitly.

Table 1 Descriptive statistics of project-level data

Campaign status	Count	Share %
Successful	7867	38.96
Failed	12,321	61.04
Total	20,188	100.00

speech, and visual information that is likely to influence the campaign perception of potential backers.

3 Data

To predict crowdfunding campaign success, we scraped Kickstarter data with a custom-build Python crawler. We restricted our data sample to projects in the categories *Technology* and *Product Design*. Additional criteria are that campaigns must include a project description, a project video, non-zero speech content, and were not canceled. The final dataset comprises 20,188 campaigns finished in the time frame from 2009 to 2017. Within this dataset, 7867 (38.96%) projects were successful by reaching their funding goal, and 12,321 (61.04%) projects were unsuccessful in meeting their goals (cf. Table 1). Our final data corpus, as outlined in Table 2, comprises 7.45 million words in text, 2.68 million spoken words, and 922,678 tags of objects in videos. The structure of the final information after text preprocessing is documented in Table 3.

For video object tags, we processed a total of 18,810 video minutes with an average runtime of 1:20 min.

3.1 Project descriptions

Project descriptions (cf. Table 3A) are mandatory information that every creator is required to provide. This

Table 2 Descriptive statistics of generated data corpus after text preprocessing

Source	Total token (words)	Vocabulary	Mean length of document
Description	7,459,121	156,109	369.48
Speech	2,683,687	55,536	132.93
Video	922,678	5417	45.70

Table 3 Data samples from *text*, *speech*, and *video* content.

(A) Text content sample

"(. . .) is a brand new robot construction system. It was designed, prototyped and engineered over the last two and a half years. (. . .) Work began in 2010 as a research project, funded by the National Science Foundation. Since then, we've been working to from a concept to a production ready robotics kit."

(B) Speech content sample

[confidence: 0.91][(...)] is a construction kit for you. Imagine to design, build and play with robots. Everything we learned from (. . .) We took all the robotic complexity inside the micro-controllers and boiled it down to elegant little blocks with simple connecting faces.

(C) Video content sample

['human, 2.03, 32.11', 'dress, 4.05, 28.10', 'tie, 5.00, 16.11', 'beard, 03.77, 18.01', 'smile, 8.89, 10.04', 'building, 16.11, 20.56', 'electronics, 26.23, 28.90', 'circular board, 26.58, 28.90', 'electronic engineering, 26.27, 28.90', 'cube, 29.03, 30.13', 'human, 31.54, 34.05', 'table, 35.91, 38.42']

textual information is in rich text form when scraped from the Kickstarter website. All project descriptions are cleaned from HTML syntax and formatted to enable their use by machine learning models subsequently. The *scikit-learn* toolkit (Pedregosa et al. 2011) is used to implement a custom tokenizer and lemmatizer for a given input text. In addition, an English stop-word list (Bird 2006) is used to remove stop words from previously lowercased text data. We further use language detection to restrict our dataset to English-speaking campaigns only, and we consider phrases of frequent co-located words, so that terms such as "new york" are not computed as separate words and hence distort the semantic context. The outlined data preprocessing is essential to reduce noise in the input data while calculating embeddings for words and paragraphs.

3.2 Speech content

As for speech transcription, we used the *Google Cloud Speech REST API*.² Using a custom-build Python script, all project video files are first transformed into mono channel *.flac audio files with *ffmpeg*. Audio files are subsequently uploaded into the Google Cloud to enable asynchronous English speech recognition via API, as a long-running operation until the end of an audio file. As for the Speech API, file URLs are used as input, while the

² <https://cloud.google.com/speech/>; accessed December 12, 2018.

output returns the *transcript* text of the speech and the average *confidence* of this transcription in a Pandas DataFrame (cf. Table 3B). The confidence value is an estimate ranging from 0 to 1, indicating how confident the Speech API is in a given transcription. A higher number indicates a greater likelihood that the recognized words are correctly transcribed. However, it cannot be guaranteed that they are correct.³ There are a few situations where some audio file transcriptions indicate a low confidence level, close to 0. In those cases, for instance, the original video either does not contain any speech signal other than music, a different language, or infrequent, untrained words. We only consider transcriptions with a confidence score of 0.80. An inspection showed that the entire body of text is sufficient, although the software does not recognize new brand (project) names and sometimes has problems with sentences that contain long stylistic pauses. With regard to preprocessing, we apply the exact same preprocessing techniques as outlined above.

3.3 Video object recognition

For visual content, we analyze all Kickstarter video files with the *Google Cloud Video Intelligence REST API*.⁴ The goal is to detect all different objects and their duration of appearance in each streaming video file (cf. Table 3). The *analyze labels* function from the Google Cloud Video Intelligence API is used to source object labels (*labels*) and their duration of appearance (*shots*) in a video sequence. In total, our data comprises 922,678 identified objects in 18,810 total video minutes with an average runtime of 1:20 min. A manual inspection of a few videos and respective video tags shows that the API has indeed a high accuracy identifying objects and events in videos. For video tags, no application of additional text cleaning was necessary.

3.4 Models

In the following, we introduce the course of the inquiry. We start with the definition of the language vector model and continue with a description of the classification methods. We then examine the results of the

classifications and utilize penalized regressions to shed more light onto predictive features in text, speech, and video content.

3.5 Language model

“You shall know a word by the company it keeps”—Firth (1968[1957]:179), cited from Jurafsky and Martin (2016).

In the past years, deep neural networks (LeCun et al. 2015) played a significant role in improving the computational models for natural language processing (NLP) and neural probabilistic language models (Bengio et al. 2003). At the core of our system is a combination of an unsupervised learning of multidimensional vector representations of words and documents, respectively, as well as a supervised labeling approach with regard to campaign outcomes. The very first challenge to process natural language using deep learning is to represent the textual data in the form of fixed-length numerical data as input for deep neural networks. The most common approaches are bag-of-words (BOW), n-grams, and one-hot vectors.⁵ However, such models either do not preserve the word order or generate the same representations for different ordered sentences with the same words. Mentioned methods maintain the short context but tend to lose the overall semantics and fail drastically, when the length of a sentence is too long (Mikolov et al. 2013b). Hence, for employing neural network language models, we use word and paragraph vectors, *Doc2Vec*, preserving the semantics of natural language information.⁶ We learn these vectors using the models as discussed by Mikolov et al. (2013b) and Le and Mikolov (2014). Paragraph vectors are an extension to Word2Vec. While Word2Vec learns to project words into a latent N-dimensional space, Doc2Vec aims at projecting a document into a latent

⁵ Goodfellow et al. (2016) provide comprehensive information on terms and methods in deep learning.

⁶ We further considered two other vectorization methods: a *count vectorization* model, which only counts term frequencies, and a *frequency-inverse document frequency (TF-idf) vectorization* model, which normalizes term frequencies across documents (Sparck Jones 1972). Our final selected model, Doc2Vec, performed slightly better among these three vectorization models across all given data sources. Distributed representations of words, as in Doc2Vec, are capable of modeling more complex relationships in data and able to preserve context and similarity encoding.

³ <https://cloud.google.com/speech/docs/basics>; accessed December 12, 2018.

⁴ <https://cloud.google.com/video-intelligence/>; accessed December 12, 2018.

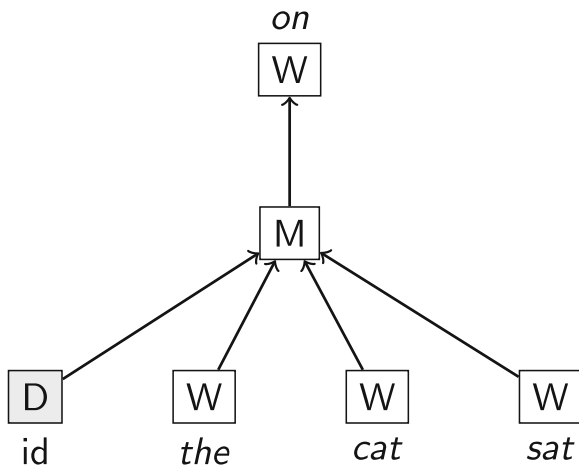


Fig. 1 Distributed memory paragraph vector model (PV-DM). The concatenation of a vector (*M*) with a context of three words is used to predict the fourth word

N-dimensional space. As such, we use Doc2Vec to learn fixed-length vector representations for each word and paragraph in a high-dimensional continuous space. More precisely, we train word vectors with a feed-forward neural network, using a bag-of-words (Fig. 1) and skip-gram (Fig. 2) approach as outlined by Le and Mikolov (2014).

Paragraph Vectors (PV) are embedding vectors which capture the overall semantic meaning of a text of variable length (“document to vector”). “*The name Paragraph Vector is to emphasize the fact that the method can be applied to variable-length pieces of texts, anything from a phrase or sentence to a large document.*” (Le and Mikolov 2014). Models of learning

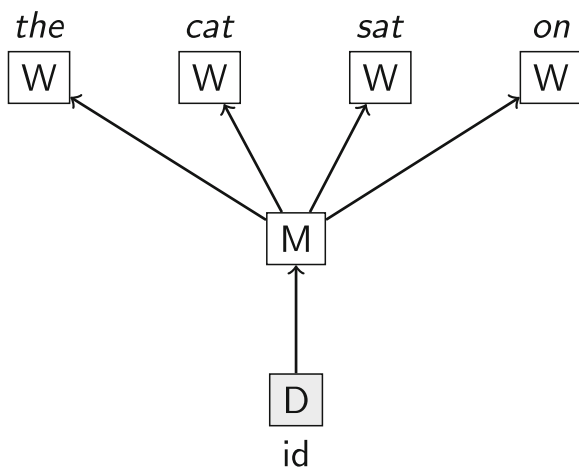


Fig. 2 Distributed bag-of-words paragraph vector model (PV-DBOW). The paragraph vector (*D*) is trained to predict a bag of context words

word vectors inspire the approach of learning paragraph vectors. According to Mikolov et al. (2013b), models using large corpora and a high number of dimensions, like the PV-DM (skip-gram) model, promise a high accuracy, both on semantic and syntactic relationships. Performance benchmarks of the Paragraph Vector approach, in comparison to other approaches such as Recursive Neural Tensor Network (RNTN) (Socher et al. 2011), Naive-Bayes Support Vector Machine (NBSVM) (Wang and Manning 2012), or Restricted Boltzmann Machines model (RBM) with bag-of-words (Dahl et al. 2012), indicate a lower error rate (Le and Mikolov 2014).

Therefore, in our implementation, we follow the suggestions of Le and Mikolov (2014) and make use of a hybrid model to generate paragraph vectors. In this model, two distinct paragraph vector models are learned as a byproduct of a classification task, where a word serves to predict its neighboring words. The number of neighboring words predicted is defined by the context window size a priori, and word embeddings are shared among all paragraphs. After being trained, the paragraph vectors are used as features for the paragraph. Eventually, initialized word vectors capture the semantic meaning of the document during the training process of a model (cf. Table 4).

We employ the following paragraph vector models: (1) Distributed Bag-of-Words (PV-DM), as shown in Fig. 1, and (2) Distributed Memory (PV-DBOW), as illustrated in Fig. 2. The PV-DM model uses the paragraph ID and given the word from a randomly sampled context window as input and *predicts all the residual words in the given context window*. The PV-DBOW model *predicts one randomly chosen word from the context window*, given the paragraph ID in combination with all the other residual words from the context window. The following analysis comprises a PV-DM and PV-DBOW model with 200 dimensions (*vector size = 200*), a word-window of four words (*window = 4*), and hierarchical softmax (*hs = 1*) (Mikolov et al. 2013a).

Video data analyses required several adjustments. The video data of each Kickstarter project contains the text labels for the objects appearing in the corresponding project’s video. These text labels are single words that define an object like “street”, “bus”, “phone”, “face”, or “computer”. Depending on the content of the video of each project, a variable number of objects are detected for each video, and hence each project contains a

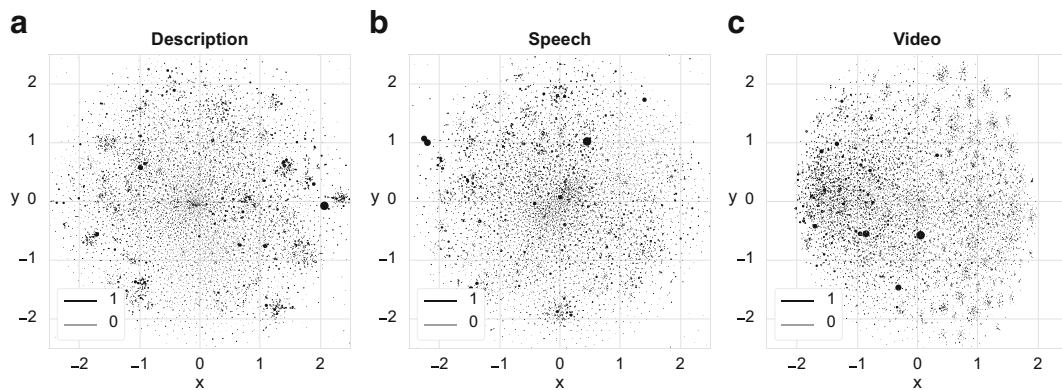


Fig. 3 Visualization of the language vector models in t-Distributed Stochastic Neighbor Embedding (*t-SNE*; Maaten and Hinton 2008). The figure shows the high-dimensional ($N_{dim} = 200$) embedding of documents clusters, downsized to a two-

dimensional system. Clusters represent documents of similar semantic meaning in the high-dimensional space. For instance, black clusters represent successful projects in semantic proximity

different number of words as text labels. As the text length in videos is naturally shorter than description or speech content and does not represent a semantic sentence structure, we decide to restrict the window-size of our Doc2Vec model to two words in order to prevent overfitting. Figure 3 illustrates a Visualization of the language vector models in t-Distributed Stochastic Neighbor Embedding (*t-SNE*; Maaten and Hinton 2008). *t-SNE* is a dimensionality reduction method that is well suited for high-dimensional data visualization. Color shades depict the document vector embedding of binary classified documents within a 200-dimensional space, reduced to two dimensions.

Table 4 provides an excerpt of the outcome of language model training that further explains the vector

representations. As the highest loading weights for the selected terms “university,” “research,” and “hardware” show, our PV-DBOW model is reasonably accurate, especially in view of a relatively small training corpus for a language model.

In order to train our two language models, we selected a range of parameters and evaluated their performance with a logistic regression as a baseline classifier. In doing so, we separate the dataset into two parts: 80% of the campaigns are selected as a training set ($N=16,150$) and the remaining 20% as a test set ($N=4038$). With regard to the hyperparameters of our paragraph vector model, we cross-validated the window size, and determine four words as the best fit for description and speech, while video models use a word-

Table 4 Most similar word vectors as measured by cosine similarity ($\cos(\theta)$) in the 200-dimensional word-vector space (PV-DBOW model)

‘university’	$\cos(\theta)$	‘research’	$\cos(\theta)$	‘hardware’	$\cos(\theta)$
state university	0.654	market research	0.580	hardware software	0.575
institute technology	0.632	study	0.573	firmware	0.474
doctoral	0.611	extensive research	0.526	software	0.465
master degree	0.594	investigation	0.495	ntp	0.448
university california	0.587	scientific	0.494	electronic	0.447
engineering university	0.579	scientific research	0.490	software hardware	0.429
college art	0.576	investigate	0.463	electronics	0.418
business administration	0.573	experiment	0.463	microcontroller	0.418
carnegie mellon	0.569	testing	0.455	hardware firmware	0.413
cornell university	0.567	experimentation	0.455	electrical component	0.409

Sample: Text description data

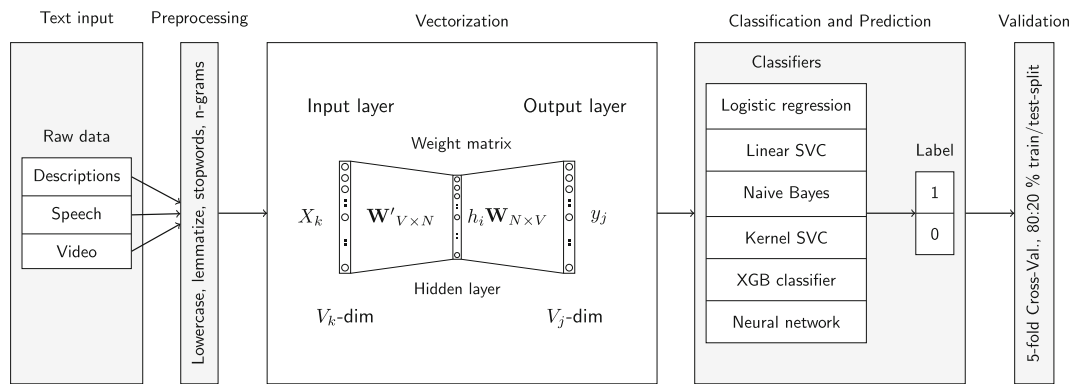


Fig. 4 Illustration of the supervised machine learning approach. Neural network example: A Continuous Bag-Of-Words (BOW) model with only one word in the context. The input of one-hot encoded vectors of multiple words is represented as X_k . The vocabulary size is V , and the hidden layer size is N_{dim} , which is the pre-defined number of computed embedding dimensions (200). The weights between the input layer and the output layer

can be represented as matrices $W_N \times V$, the input word matrix, and $W^T V \times N$, the N -dimensional embedding vector. h_i is the vector representation of a given input word. Using $W^T V \times N$, the probability score y_j for each word in the vocabulary is computed. In summary, the *Output* layer is the fixed-length vector representation of variable-length documents.

window of 3. Every subsequent language model, which inputs each the full dictionary of unique words, is computed in 200 vector dimensions (N_{dim}) (Fig. 4).⁷

3.6 Classification

After training the paragraph vectors, the 200-dimensional features are fed into several distinct classifiers. In total, six widely used parametric and non-parametric classifiers are being applied. As linear classifiers, we consider a (1) Logistic Regression and a (2) Linear Support Vector Classification (LinearSVC). As non-linear classifiers, we use a (3) Gaussian Naive Bayes (GaussianNB), (4) Support Vector Classifier (SVC) with a radial basis function kernel (rbf), the (5) XGBoost (XGBoost), which is a scalable tree boosting system (Chen and Guestrin 2016), and a (6) Multi-Layer Perceptron (NeuralNetwork), which is a neural network model with 100 hidden layers and a rectified linear unit (ReLU) activation function (Nair and Hinton 2010). As

it concerns the parameters of our classifiers, we train our classification model with Grid-Search, supported five-fold cross-validation, and iterate over a comprehensive set of individual hyperparameters in scikit-learn (Pedregosa et al. 2011). The final results represent the outcome of each best-selected classifier, by accuracy. For an overview and explanation of the related classifiers, we refer to Varian (2014), Mullainathan and Spiess (2017), and Puranam et al. (2018), who discuss in detail several machine learning classifiers widely used in the economic sciences.⁸

Crowdfunding success is implemented as a binary variable indicating whether the campaign reached the funding goal (1) or not (0). This binary representation also resembles the “All-or-Nothing” (AON) approach of Kickstarter (Cumming et al. 2015). The AON model involves the entrepreneurial firm setting a fundraising goal and keeping nothing unless the goal is achieved. Each classifier is trained using the transformed paragraph vectors as the features (inputs) and labels as outputs.

Our full workflow is implemented using the gensim (Řehůřek and Sojka 2010) and scikit-learn (Pedregosa et al. 2011) libraries in Python.

⁷ For the PV-DM model, we further concatenate context vectors ($dm_{concat} = 1$), while the PV-DBOW is set to train word vectors simultaneous with DBOW doc vectors ($dbow_{words} = 1$). Both models only consider words with a total frequency of 5 ($min_{count} = 5$). In total, we train each model with 20 iterations ($epochs = 20$) over the corpus. After applying matrix optimizations via hierarchical softmax, the description text, speech, and video models compute with an input dictionary (V in Fig. 4) of 45,670, 19,110, and 3796 unique tokens, respectively. For further reference on model parameters, see the gensim and scikit-learn documentation (Řehůřek and Sojka 2010; Pedregosa et al. 2011).

⁸ In particular, appendix 1 of Puranam et al. (2018) and Pedregosa et al. (2011) provide a good overview on the functional forms, loss functions, and regularization techniques.

Table 5 Comparison of classifiers

Classifier/data	Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1
Stratified baseline	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
Description								
LogisticRegression	0.71	0.72	0.71	0.71	0.71	0.71	0.71	0.71
LinearSVC	0.71	0.72	0.71	0.72	0.71	0.70	0.71	0.70
GaussianNB	0.64	0.64	0.64	0.58	0.61	0.63	0.61	0.62
SVC	0.67	0.70	0.67	0.61	0.71	0.72	0.71	0.71
XGBoost	0.69	0.70	0.69	0.67	0.70	0.70	0.70	0.70
NeuralNetwork	0.70	0.70	0.70	0.70	0.66	0.66	0.66	0.66
Speech								
LogisticRegression	0.66	0.67	0.66	0.67	0.65	0.66	0.65	0.65
LinearSVC	0.67	0.67	0.67	0.67	0.65	0.66	0.65	0.66
GaussianNB	0.66	0.65	0.66	0.65	0.64	0.64	0.64	0.64
SVC	0.68	0.67	0.68	0.67	0.66	0.66	0.66	0.66
XGBoost	0.67	0.66	0.67	0.64	0.66	0.65	0.66	0.64
NeuralNetwork	0.62	0.62	0.62	0.62	0.61	0.61	0.61	0.61
Video								
LogisticRegression	0.65	0.65	0.65	0.65	0.63	0.63	0.63	0.63
LinearSVC	0.64	0.65	0.64	0.64	0.64	0.63	0.64	0.63
GaussianNB	0.60	0.61	0.60	0.60	0.60	0.60	0.60	0.60
SVC	0.66	0.65	0.66	0.65	0.64	0.63	0.64	0.64
XGBoost	0.66	0.65	0.66	0.65	0.64	0.62	0.64	0.61
NeuralNetwork	0.65	0.65	0.65	0.65	0.65	0.63	0.65	0.61

Accuracy (Acc.) is the number of correct predictions, divided by the total number of predictions made. *Precision* (Prec.) is also referred to as the positive predictive value, while *Recall* (Rec.) is the true positive rate, or sensitivity. The F1-score (F1) is the harmonic mean of precision and recall ($F_{0.5}$). Values represent the weighted average of each classification

4 Results

With classification accuracy ranging individually from 60% to 72%, our results in Table 5 suggest that the textual descriptions of project creators, the words they speak, and the objects they show in videos can help to predict the outcome of a campaign. Overall, this accuracy is a 10% to 20% improvement over a stratified baseline model that is an a priori probability calculation, based on the distributions as provided in Table 1. Both Logistic Regression (LR) and a Linear Support Vector Classifier (LinearSVC) exhibit the best classification, which suggests that the classification of campaign success might be determined by partially linearly scaled features in our data. Non-linear classifiers confirm the obtained results, albeit with minimally lower accuracy. In general, the PV-DBOW model performs slightly better as compared to a PV-DM model. Despite the marginality, this finding is well in line with recent

empirical assessments of the Doc2Vec model (Lau and Baldwin 2016). The outcome classification is most accurate for predictions on description text, with an accuracy of about 71%, followed by speech with about 67% accuracy. Predictions with video content show the lowest accuracy with about 65%. Yet, despite the average length of only 60 tags, and considering the sparsity of this information, the accuracy of video tags is still surprising.

In order to further elaborate on the accuracy of the model, we inspect the example of a Logistic Regression (LR) Classification in Table 6. Logistic regression did not only perform as each one of the best two classification models but it is also a well-interpretable algorithm that is used in subsequent penalized feature analyses in this paper.

For explanation, *Recall* (or sensitivity) indicates the true positive rate, the proportion of successful campaigns that were correctly predicted as such by the

Table 6 Classification report of Logistic Regression vs. Stratified Baseline—model: PV-DBOW

Data/class	Support	Prec.	Rec.	F1	F1- Δ
Stratified baseline					
0—unsuccessful	2396	0.60	0.60	0.60	–
1—successful	1642	0.38	0.38	0.38	–
Weighted mean	4038	0.51	0.51	0.51	–
Description					
0—unsuccessful	2396	0.73	0.85	0.78	+18%
1—successful	1642	0.71	0.54	0.61	+23%
Weighted mean	4038	0.72	0.72	0.71	+20%
Speech					
0—unsuccessful	2396	0.68	0.83	0.75	+15%
1—successful	1642	0.64	0.44	0.52	+14%
Weighted mean	4038	0.66	0.67	0.66	+15%
Video					
0—unsuccessful	2396	0.66	0.84	0.74	+14%
1—successful	1642	0.62	0.36	0.46	+8%
Weighted mean	4038	0.64	0.65	0.63	+12%

model. *Precision* indicates the proportion of positive results that are true positive results. A lower *Precision* score is indicative for a high prevalence of false positives (“Type-I errors”). The *F1-Score* is a harmonic mean of Precision and Recall. As Table 6 shows, the F1-score against a baseline model improves by an absolute 18% for non-successful campaigns (“0”) and by about 23% for successful campaigns (“1”), using description data only. In-class predictions for speech yield similar results, with an overall prediction improvement of about 15% on average. Video tags improve the prediction by an absolute 12% for each class. We infer from this that there may be greater variance within the video data (cf. Fig. 3) and that “vectors of success” are hence more difficult to classify. However, in light of the nature of the data source, on average about 1:20 min of video, it seems remarkable that the classifier correctly identifies on average about 74% (F1) of unsuccessful campaigns.

Figure 5 shows confusion matrices of each classified data source. The confusion matrix describes the classifier’s performance on a set of (out-of-sample) test data for which the “true” values are known. Overall, the current algorithm is better suited to find campaigns that are likely to fail. As for the case of Recall in speech data, the model identified about 83% (1978) out of 2396 unsuccessful campaigns in the test set correctly as

unsuccessful. For successful campaigns, the algorithm classifies 44% (728) out of true 1,642 successful campaigns correctly as successful. For all data sources, unsuccessful campaigns are better predicted than successful campaigns, even after considering class weight adjustments in the classifiers’ parameters. Across the three data sources description text, speech, and video content, and as measured by F1-score, the classifier is an absolute 20%, 15%, and 12% better in identifying non-successful campaigns than successful campaigns. Overall, we conclude the results to be robust and in line with scores reported in previous studies on predictions with text data (Greenberg et al. 2013; Mitra and Gilbert 2014; Du et al. 2015).

Worth mentioning, in a deeper inspection of learning curves of our models, we find that the test accuracy of our model asymptotically approaches 72.0% at a training set size of about 10,000 projects already. This indicates that the used 16,150 projects seem to be a sufficient training set size for our model.⁹ However, despite only a minor improvement of accuracy, we would still expect that more training data will improve the predictive accuracy, for instance due to lower variance within the data.

4.1 Feature union

In order to evaluate the predictive power of all of the information sources combined, we follow the procedure as outlined in Section 2. After training the models, we concatenate the respective composed feature columns into one new feature matrix. The new matrix, a feature union, is then trained with a Logistic Regression. Applying a fivefold cross-validation, we train the statistical model based on 80% training data and then apply the learned estimator to 20% test data.

As Table 7 shows, the most accurate model (73%) is *M4*, with “text, speech, and video” data combined (73% F1-score), on par with combinations of *M1* “text and speech” and *M2* “text and video” information. Overall, feature union improved the prediction by an absolute 2%, as compared to the best single-source prediction with description text only (see Table 5: 71% vs. Table 7: 73%). Using speech and video information, model *M3* achieves a 67% F1-score. While the accuracy for speech and video content only does not seem high, despite a 16% improvement over the baseline, achieving the prediction score (and a 84% Recall for unsuccessful campaigns)

⁹ $20,188 \times 0.80$ (training size).

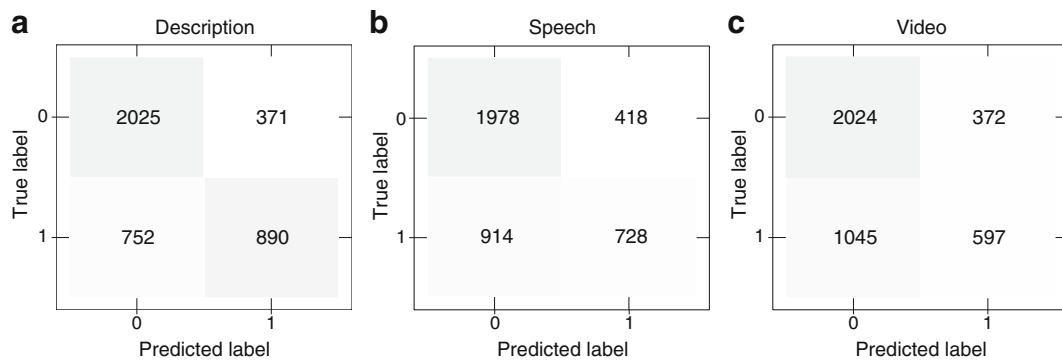


Fig. 5 Confusion matrices for *text*, *speech*, and *video* classification—model: PV-DBOW, logistic regression classifier

with only 1:20 min video time on average is still a surprising result. We would expect video data to provide an even better contribution, once additional features, such as labeled events or emotional arcs of storytelling, are included.

In feature union regressions, information provided in videos is more unambiguous, as the β -coefficients suggest. Here, we conclude that in model *M4*, unifying text, speech, and video information, video features contribute most to predictions with regard to their relative power (0.62), as compared to description text (0.23) and speech information (0.15). This pattern matches to results for model *M2* and *M3*. With regard to similarities of prediction accuracy among the different models, however, we conclude that all information sources combined potentially comprise similar information and hence do not significantly improve the predictions in feature union models, as compared to single-source predictions (Table 5). Controlling for the influence of the volume of information provided in each data source, model *M5* indicates a significant influence of length of description, speech and video token on the prediction of outcomes.¹⁰

Our conclusion is that in particular, video information improves a combined prediction; however, similar F1-scores among models suggest that the entropy of information across features seems rather complementary. Hence, we argue that different features mostly seem to underpin another predictor's results.

¹⁰ This result resonates with a Pearson correlation analysis of feature lengths and project outcomes. More description text and more tags in videos correlate positively with funding success ($p \leq 0.01$), while the length of speech token is insignificant.

4.2 Predictive words

One of the most intriguing features of machine learning algorithms is the ability to generate stylized facts that can induce explicit quantitative inductive inferences (Puranam et al. 2018). What is interesting in this work is that we can further corroborate previous effects found for the linguistic style. Yet, these were often thought to relate primarily to pro-social businesses (Parhankangas and Renko 2017). At the contrary, garnering social support in commercial crowdfunding also places a strong emphasis on higher order motivations rather than monetary contributions.

In Fig. 6, we capture the importance of words with respect to other documents in a corpus, as classified by a penalized logistic regression against binary labels. By doing so, we try to open the “black box” of the machine learning algorithm, i.e., we try to infer the potential weights in the hidden-layer network of our PV-DM and PV-DBOW neural network model.

When we investigate the most predictive words within the different textual, linguistic, and visual representations, we find that all terms related to monetary depictions of the venture reduce the chances to reach the campaign goal successfully (cf. Mitra and Gilbert 2014; Kaminski et al. 2017). In Fig. 6, we show each the top 25 predictive terms for a successful (“1,” black) and unsuccessful (“0,” gray) outcome.

For textual descriptions (reported in Fig. 6a), legitimizing activities that are often thought to help a venture connect to external stakeholders such as patents, prototypes, or money are among the worst textual descriptions to be used. At the contrary, and as already found in Mitra and Gilbert (2014), linguistic styles in text content that aim to trigger excitement (“amazing”), social (“backer,” “community,” “thank”), or technical inclusiveness (“open

Table 7 Classification results of combined data sources

Model	Acc.	Prec.	Rec.	F1	D	S	V	F
M1: D + S	0.73	0.72	0.73	0.72	0.60	0.40	–	–
M2: D + V	0.73	0.72	0.73	0.72	0.32	–	0.68	–
M3: S + V	0.68	0.67	0.68	0.67	–	0.29	0.71	–
M4: D + S + V	0.73	0.72	0.73	0.73*	0.23	0.15	0.62	–
M5: D + S + V + F	0.72	0.72	0.72	0.72	0.05	0.04	0.13	0.78

Classifier with the highest accuracy is marked with an asterisk (*). D=description, S=speech, V=video, F=additional text features: length of description, speech, and video token. Feature importance is measured as the standardized mean of all absolute β -coefficient values

source”) are better predictors of campaign success than firm-level determinants. Figure 6a also reports that indications of early-stage developments (such as “idea,” “prototype,” or “concept”) are negative predictors. This finding may hint at the riskiness of the campaign as perceived by the potential backer. In contrast, successful campaigns report signals related to “press,” “update,” and “stretch goal,” indicative for more maturity and future goals. Likewise, speech patterns (Fig. 6b) that concentrate on high-order meanings (such as “perfect,” “amazing,” “excite,” “super”) carry high weights in explaining success. In addition, words indicating distinct product features such as “tiny,” “titanium,” “python,” “super easy,” “arduino,” “compatible,” and “compact” are also positive predictive terms. Lastly, animations, cartoons, illustrations, photo-montages, special effects, or depictions shown in videos have negative consequences for campaign success. It appears as if potential backers are more interested in people (“team,” “student”), and products or product demonstrations (“experiment,” “laboratory,” not shown—“3D Printing,” $\beta + 0.55$) rather than sketches thereof (as can be seen in Fig. 6c). Even more, one may see “street-credibility” in the positive β -coefficients of objects in successful campaigns (“passenger,” “street,” “backpack”). Tools and accessories shown in videos (“office supplies,” “pen,” “electronics accessory”) also have a positive influence on reaching the campaign goal.¹¹

5 Discussion

Because early-stage product financing is often more difficult to secure for new firms due to information

asymmetries and other liabilities of newness, an entrepreneur must find ways to meet the expectations of various audiences with differing norms, standards, and values as the venture evolves and grows from the conception stage to potential commercialization (Fisher et al. 2016). Therefore, the act of crowdfunding by entrepreneurs may be seen as means to gain strategic legitimacy, as the entrepreneur looks to purposefully manipulate and deploy symbols in order to garner positive legitimacy judgments (Suchman 1995). Work in the entrepreneurial finance literature has already emphasized the role of visual cues of financiers’ decision-making (Chan and Park 2015). Similar to startup pitches to venture capital investors, or business angels, potential crowdfunding investors will underlie time constraints. There is virtually no way to compare the product offering seen in a campaign with other potential products one might be interested in. Hence, potential backers will have to rely on shortcuts or heuristics to make decisions.

This also emphasizes the need to understand crowdfunding from a consumer’s perspective. Most studies of new venture development take an entrepreneur/firm perspective to understand how firms are created and novel products are brought to market. Yet to understand value appropriation in an early stage market, a consumer perspective might be key. Much has been written about the need to employ minimum viable products and to engage customers into the development of new products (Blank 2013; Ernst et al. 2010). Potential consumers in crowdfunding campaigns perceive new products and ultimately decide the fate of the new product development process. Importantly, in crowdfunding, potential consumers employ a taste-based approach (Chan and Parhankangas 2017). This underscores the need to align potential consumer’s expectations with the perception of entrepreneurs and the products they pitch in their campaigns.

¹¹ Note for video information that “ammunition” and “knife” hint at the possibility of false positives in the item identification (“algorithmic bias”). In manual inspection, we found that these item identifications related to sequences with tools and equipment from the workbench.

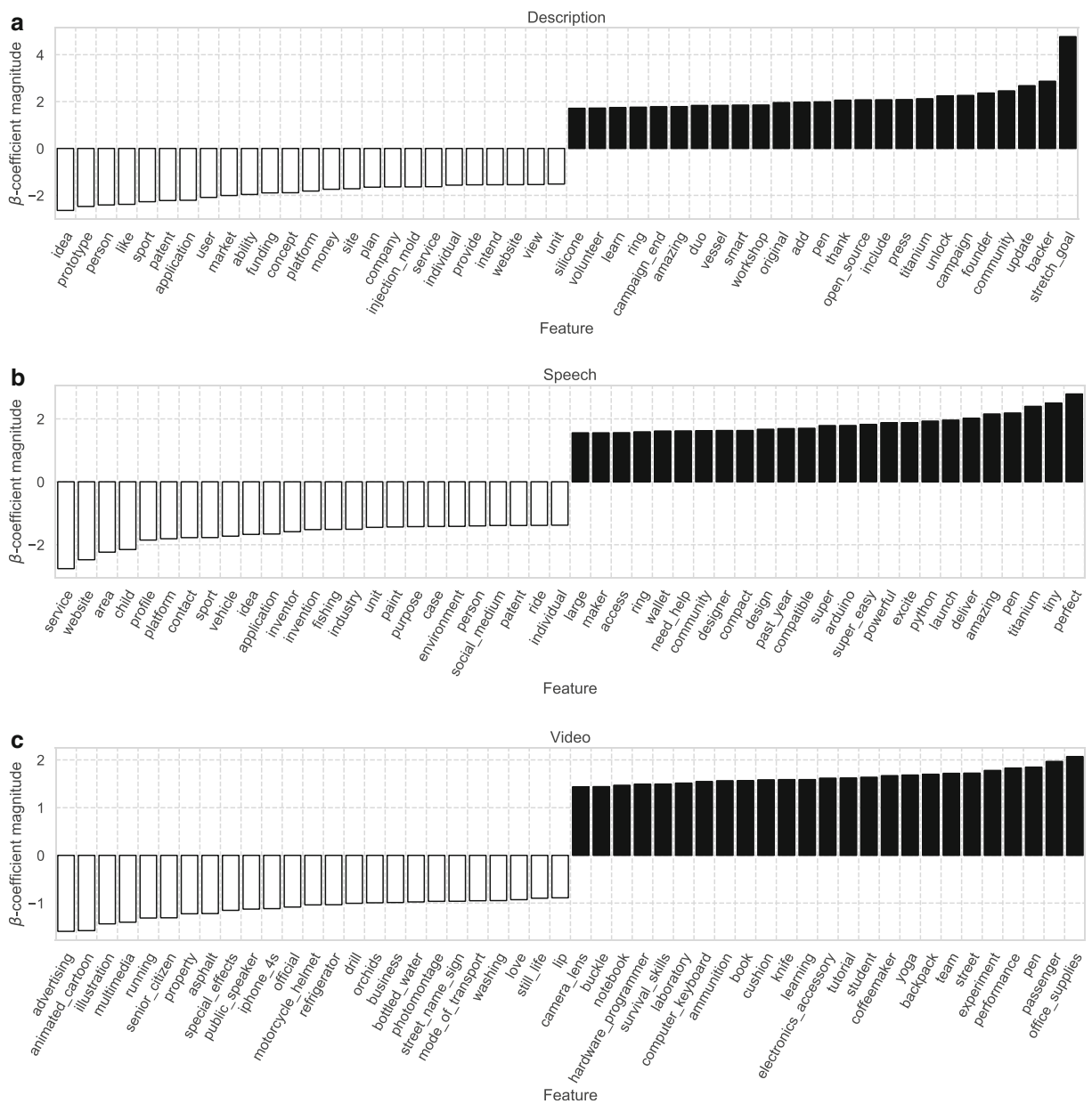


Fig. 6 Tf-idf regression β -coefficients for highest scoring features. Based on training set ($N = 16,150$)

This study therefore employs a neural network and natural language processing approach to predict the outcome of crowdfunding startup pitches using text, speech, and video object-related metadata in 20,188 crowdfunding campaigns. While prior work has predominantly studied a single aspect of communication in isolation such as impression management (Parhankangas and Ehrlich 2014), competence signaling (Gafni et al. 2019), or persuasion (Allison et al.

2017), we can combine textual, visual, and language information to provide a more complete picture of the communication process between the consumer and the entrepreneur. Consequently, the approach and method explained and applied in this work may help to guide theoretical researchers in focusing on theories that best explain the phenomenon they are interested in (von Krogh, 2018). Machine learning approaches may help to take stock of theoretical explanations. Researchers

will find identical (or very, very similar) conclusions when observing the same dataset or studying another large crowdfunding dataset. Hence, findings are robust and generalizable. Even more, the pre-trained vectors of our language models in description, speech, and video content could be used in similar contexts (“transfer learning”). This, in essence, may help us to strive for parsimony and avoid the superfluous.

Figure 6 helps to delineate the different textual and visual cues present in a campaign that ultimately coalesce in the potential consumer’s decision-making processes. By examining specific textual and visual predictors, we can account for connections, similarities, and complementarities between signals and cues presented.

Linguistic styles were often thought to relate primarily to pro-social businesses. Yet garnering social support in commercial crowdfunding also places a strong emphasis on higher order motivations rather than monetary contributions. When we investigate the most predictive words within the different textual, speech, and video representations, we find that all terms related to monetary depictions of the venture reduce the chances to successfully reach the campaign goal (Mitra and Gilbert 2014; Kaminski et al. 2017). Hence, non-monetary motivations are important cues in reward-based crowdfunding. For textual descriptions (reported in Fig. 6a), legitimizing activities that are often thought to help a venture connect to external stakeholders (such as patents or outward marketing) are among the worst textual descriptions to be used. At the contrary, linguistic styles that aim to trigger excitement (perfect or amazing) or are aimed at inclusiveness (you, community) are better predictors of campaign success than firm-level determinants. These patterns hold for speech information.

Figure 6b also reports that indications of early stage developments (such as prototype or concept) are negative predictors. This may hint at the riskiness of the campaign as perceived by the potential backer. The perception of uncertainty as it relates to uncertainty about the technical feasibility due to the early product stage of the campaign may therefore be detrimental for campaign success. This is important, as Stanko and Henard (2017) document that crowdfunding campaigns have on average only completed about 60% of activities such as developing the product’s feature set, conducting business analysis, prototyping, engineering/design/coding, etc. This adds substantial uncertainty about whether or not crowdfunding campaigns can actually live up to

rosy expectations. Similarly, prior work has shown that campaigns that propose radically different solutions may reduce the chances of the campaign to reach its funding goal (Chan and Parhankangas 2017). Uncertainty perceptions may substantially reduce evaluations of new products and reduce purchasing intentions among potential funders (Biswas and Biswas 2004).

Along these lines, our results also report that the time of possible product possession may have a detrimental effect. Illustrations or depictions shown in videos have negative consequences for campaign success. It appears as if potential backers are more interested in products rather than in sketches thereof (as can be seen in Fig. 6c). This also highlights the necessity for crowdfunding campaigns to find ways to overcome perceptions of uncertainty. Allison et al. (2017) report that crowdfunding campaigns often omit details or schematics of the proposed technology, likely due to potential risks of imitation by competitors. Instead of referring to prototypes absent detailed information, crowdfunding campaigns could employ peripheral cues to illustrate the benefits of their product to increase awareness and reduce perceptions of uncertainty. The marketing literature has shown that analogies in use of a new product may help to overcome negative product evaluations (Goode et al. 2010). Similarly, drawing attention as to why potential consumers would benefit from funding the campaigns may also overcome perceptions of uncertainty (Castano et al. 2008).

While prior work by Parhankangas and Renko (2017) argued that commercial entrepreneurs need to primarily focus on product, or firm and entrepreneur-related signals, our findings highlight signals that make the campaign more emotionally appealing and cognitive salient to predict campaign success best (Allison et al. 2017; Parhankangas and Renko 2017). Our work also highlights potential areas for future theorizing. Work in psychology and marketing has emphasized that the environment in which signals are sent has a profound impact on how information is construed by receptors of said signal. If psychological distance is high (let that be temporal, spatial, and social) between an individual and an object (a new product concept, for example), higher level abstractions (likely omitting secondary or peripheral information) work best in increasing receptivity, a consumers’ conscious (or unconscious) willingness to react positively to a signal received (Dhar and Kim 2007; Trope and Liberman 2003). The results show that speech patterns that concentrate on high-order meanings (such as

beautiful, super, amazing) carry high weights in explaining success. These results are important, as positive psychological language is a “*costless signal*” (Anglin et al. 2018). Our findings emphasize that positive psychological language is salient in environments where objective information is scarce and where investment preferences are by and large taste based.

When it comes to explaining crowdfunding success, textual descriptions (written text) and visual representations matter. Prior work reporting the importance of the benefit of linguistic patterns for campaign success was only derived in the absence of visual information (Parhankangas and Renko 2017). Consequently, the accentuation of product benefits has implications for how information is construed by consumers (Lieberman and Trope 1998). It appears that the sight and vision of this information (reading and watching) is similarly important. During the emergence phase of a new venture, visual communication appeals to the target audience. Similarly, our findings also attest to the role of quickness and speed in allowing consumers to make decisions. On the venture level, there is mounting evidence that quicker is not always better, and that ventures should take time to organize their venture activities (Kim et al. 2015; Brush et al. 2008). To the contrary, our results in here show that when it comes to designing first impressions for customers, it is important to allow for quick and fast impression of the product and the benefits it may bring.

Prior work has found exploratory evidence that positive psychological language in video transcriptions does not affect crowdfunding performance (Anglin et al. 2018). However, our results report a hierarchy between the different media employed. As such, the state of attention matters as to how different media embedding are to be evaluated. In low attention states, potential backers might be more responsive to cues, such as appealing and attractive graphics or an overly enthusiastic language or presentation (Allison et al. 2017). Videos are shown at the top of the crowdfunding page. Employing enthusiastic language or showing the product in action may capture potential backer’s attention. Only if the video induces a high attention state will individuals be willing to evaluate subsequently shown textual material, narratives, or schematics in more detail. Higher-quality videos lead potential consumers to form positive impressions about the entrepreneur and the campaign and may elevate the perception of other signals provided such as textual descriptions (Scheaf et al. 2018).

We show that visual, potentially emotionally appealing cues are the most potent signals crowdfunding campaigns can provide. Absent of inducing high attention states, written text often fulfill a ceremonial role only, where entrepreneurs show that they conform to expectations (Honig and Karlsson 2004). Business plans, for example, are often evaluated to have the right length, form, or document structure. Written text therefore is mostly ceremonial; it does reveal that the individual understands the rules of the game (Kirsch et al. 2009). However, the most important signal are information that are not easily inferable from written text and often specific to a given product or business opportunity and that capture the attention of potential backers. In videos, potential campaigns are more likely to be receptive if they see the product in action, rather than sketches thereof or in an unappealing context. Obviously, linguistic expressions in text and speech that are abstract and more emotionally salient work better in increasing campaign success.

This also opens the door for more experimental approaches to better understand how visual aesthetics affect crowdfunding campaign success. How can subtle changes to visual context better transmit the message of the crowdfunding campaign and increase receptivity among potential backers? Body language, ambience, tone, and voices may all affect how potential backers react to crowdfunding campaigns. Hence, we believe that the notion of how information are visually construed in online campaigns is an area that warrants further attention for both theory building and empirical inferences. From a practical perspective, it becomes important to effectively communicate and present oneself and the product (Gafni et al. 2019). An entrepreneurial appearance that suggests creativity and passion may increase the chances to successfully pitch for capital contributions (Davis et al. 2017).

6 Implications

“The tendency of these new machines is to replace human judgment on all levels but a fairly high one, rather than to replace human energy and power by machine energy and power.”—Norbert Wiener, 1949

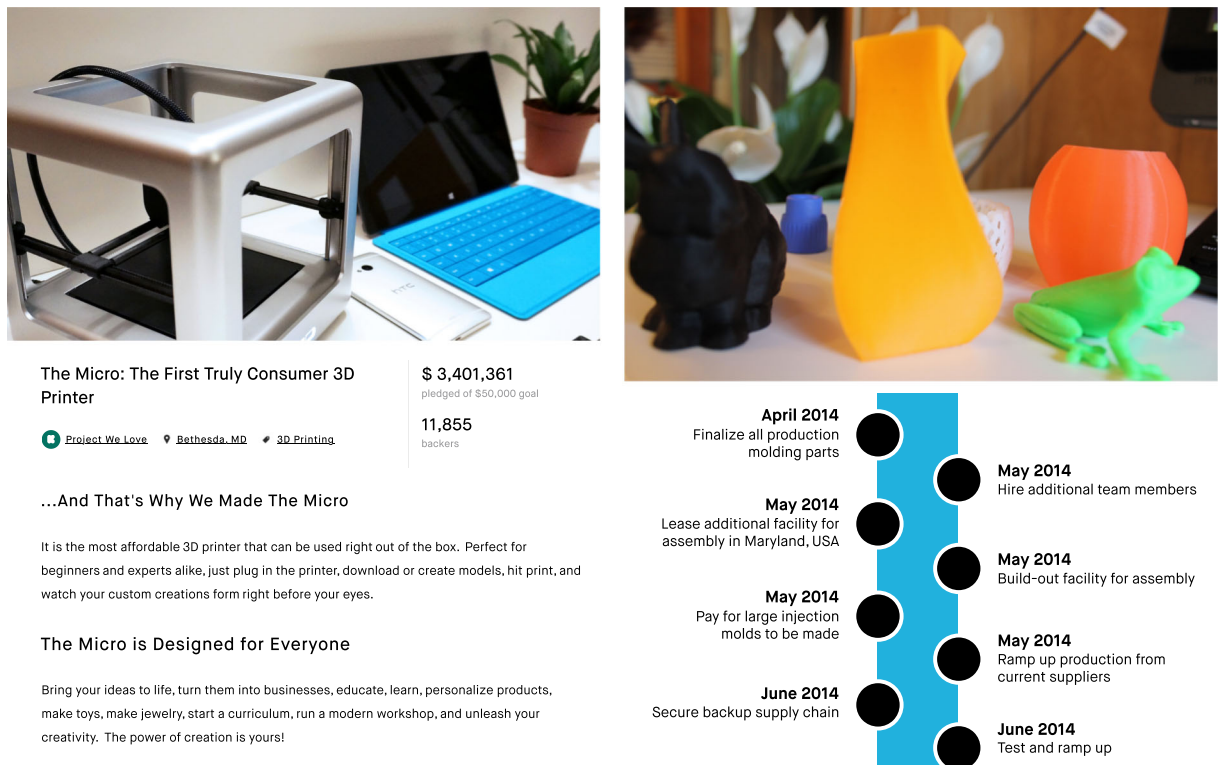


Fig. 7 “The Micro: The First Truly Consumer 3D Printer”—The pictures showcase the product’s look and feel, an explanation of benefits right below the video, as well as potential outcomes and a campaign timeline.

To conclude with an example, the *Micro* (<https://www.kickstarter.com/projects/m3d/the-micro-the-first-truly-consumer-3d-printer>) is a consumer 3D printer from Bethesda, MA, launched on Kickstarter in April 2014 (Fig. 7). The campaign had the goal to raise \$50,000 from the crowd and eventually crossed the bar of more than \$3.4 million, contributed by more than 11,800 backers. In as little as 25 hours, the campaign had raised over a million dollars already. The campaign highlights prominently the various insights derived from our empirical analysis: To begin with, visual information on the top of the page shows the product in action. Instead of sketches or still images, the potential consumers can directly get a “look and feel” on the campaign page images and video, and learn more about the people behind the product. As it relates to language employed, the campaign uses inclusive language that emphasizes why consumers should buy and support the product, and not on how the product is being used: “*The Micro is Designed For Everyone—Bring your ideas to life, turn them into businesses, educate, learn, personalize products, make toys, make jewelry, start a curriculum, run a modern*

workshop, and unleash your creativity. The power of creation is yours.” The linguistic style employs words such as “fantastic; enjoyable; fun.” Technical specifications and product features are only introduced after consumers have been set into a state of high attention and after the product has been made emotionally appealing and salient. The campaign page also includes a tentative timeline to reduce uncertainty about product development and delivery. Again, this information is shown after a state of high attention has been induced. Overall, this example of a successfully funded product (among others) fits well to predictive features as shown in Section 3.5.

However, the practical implications of this study allow us to look beyond crowdfunding. As startups nowadays can be launched in the crowd (i.e., Kickstarter or Indiegogo), cloud (i.e., Stripe Atlas, AngelList), and as startup programs invite entrepreneurs to pitch via text and videos online (i.e., Y Combinator), the proposed method can be useful for structuring and sorting available information. In another perspective, our approach is using crowdsourced labels on viable business ideas, combining human and computer intelligence. As such, information

embedded in textual information and funding success is a continuous process of semi-supervised learning on business ideas, an example of a *human-in-the-loop* machine learning system. For the long term, we expect a more advanced system that mimics the human capabilities of visual computation in investment decisions. Such a system has the power to apply artificial intelligence to the evaluation and feedback process of presented venture ideas.

Illuminating the complex nature of new venture organizing offers several academic as well as practical implications. In VC financing, firms are 23 months old when they obtain funding (Kaplan et al. 2009).

Hence, there are many opportunities to employ machine-learning techniques to predict the potential for firms applying for VC funds. This is especially important, as Kerr et al. (2014) report low correlations between investor assessments and future funded firm success. Even more worrisome, VCs often take a pass on later successful investments. Improving predictions based on communication characteristics of startups may allow to better select candidates for VC funding, especially considering that even experienced investors have a “bounded rationality” (Simon 1955) and that the venture market is a market with potential “lemons” (Akerlof 1970).

7 Limitations and future research

Combining metadata of information in product pitches, we propose a machine-learning approach to train a vector language model and a logistic classifier in identifying the features of successful and non-successful entrepreneurs. The very novelty of this contribution is the volume of the dataset, comprising not only description but also speech and video-related data.

Applying novel techniques such as machine learning comes with potential caveats that may point to future research and areas of potential improvement. First, our data does not rely on convenience samples used to study a priori-derived hypotheses. Rather, we ask more audacious research questions and explore the data patterns in light of this task. Our research provides a substantial basis for inductive theory building, but cannot provide evidence for or against previously derived hypotheses.

The patterns detected are robust to spuriousness caused by omitted variables that would imply alternative explanations to our findings. The neural network

accounts for a multitude of alternative combinations of variables and higher order interactions. The main tenet that affects the robustness and replicability, however, is over-fitting. Our model therefore goes long ways in assessing the validity and robustness of the estimates to allow meaningful interpretations. In this light, high-quality data and diligent analyses are of outmost importance to make our findings generalizable and replicable. We rely on several validations; training and prediction data subsets to ensure that we can infer causality from the underlying prediction derived.

The defining parameter of “big data” is the fine-grained nature of the data itself, thereby shifting the focus away from the number of participants to the granular information about the individual (George et al. 2014, 321). Instead of eliciting responses from consumers, we can directly predict human behavior based on the response to communicative stimuli in crowdfunding campaigns. Yet, the underlying motivation for the respective individual that contributes to the campaign remains unobservable. It is therefore important not to forego alternative data collection mechanisms that can further help to delve into the reason why we observe the patterns we observe. Is it because the product was deemed extremely useful? Was the product technologically advance and geeky? Did the reward structure offered fit the product benefits presented? While open data in crowdfunding helps to gain insights into the broader factors at work, it is, in our view, important not to leave the micro mechanisms at play out of sight. Along these lines, moderators of relationships studied in here are equally important to derive practical implications for different entrepreneurs with different degrees of innovative new products in different stages of development.

Even though our model achieves a moderate accuracy in predictions with video data, the models can likely be improved by employing more granular and accurate data from visual computing API. For description data, available information might be enriched by applying text recognition (OCR) on images on campaign websites. Most Kickstarter campaigns come with a variety of images in the description text, containing important information. The prediction accuracy in speech data has the potential to be improved by more accurate transcriptions, or alternative APIs.

One further limitation relates to the nature of our data. Crowdfunding is a very special investment setting, where the investment ratio of funders may not always follow

financial motives. Investors seek unique solutions to problems they encounter or gadgets they may perceive as attractive (Gerber et al. 2012). Though past research has shown that crowdfunders show expertise in investment decisions, just like professional venture capital investors (Mollick 2013; Mollick and Nanda 2015), future research might consider the limitations of crowdfunding as a training dataset for identifying features of successful entrepreneurs, for instance in VC startup investments.

Aside from shortcomings in the comparability of VC and CF startups, we may have also overlooked the fact that the machine can embed human error and algorithmic biases in the learning process. If for any reason, a crowdfunding project raises more than \$1 million to create wristwatches for dogs, the algorithm will learn that such a product might be a “good” idea. As such, the label “successful funding” is very constrained in being a marker of business viability. Likewise, the successful funding has limited explanatory power with regard to the actual entrepreneurial results. A campaign can be over-funded but nevertheless entrepreneurially unsuccessful by failing to deliver products (Mollick 2015). Therefore, future explorations should consider to extend the label of “success” toward measurable economic results such as Amazon listings, an active website, or other measures of a product’s market performance after crowdfunding (cf. Stanko and Henard 2017).

Future work on a “*prediction machine*” (Agrawal et al. 2018) may also consider the differentiation of features that relate to *people* or *product*. Professional investors often highlight that people are more important than ideas: As Ron Conway, an experienced angel investor, states: “*Well, we invest in people. I’ve been doing this for 21 years, and I have talked to thousands of entrepreneurs. I’m not looking at their idea. I’m looking at: Are they a leader? Are they focused on their product? Can they attract the team? What are the co-founders like? I can tell within three minutes.*” (Chafkin 2015). In this regard, founder characteristics, in particular from speech and video content, may already be implicitly represented in the paragraph vector space.

Likewise the *ImageNet* (Deng et al. 2009) and Large Scale Visual Recognition Competition (ILSVRC) in visual computing, a similar challenge for predicting promising startup projects could be of interest for entrepreneurship research. Prospective research might also consider a more controlled experimental setting in which the investment rationale of professional venture capital investors is being compared to a Bayesian

decision-making system, as outlined. Towards the other end of this thought, we could also ask: Can machine learning improve investors’ (or entrepreneurs’) decision-making through feedback with data? Therefore, it could also be examined in real-world experiments, how human decisions can be improved by a supportive machine. This approach would not only evaluate the precision of the systems but also help to analyze the benefits of information systems in human–computer interaction (Licklider 1960). Using artificial intelligence, we can potentially augment human intelligence in innovation investment decisions and enable “*cyber-human learning loops*” (Malone 2018, 234).

8 Conclusion

Recent work in crowdfunding research has focused on observational studies that allow for causal interpretations of data by adjusting for observed differences in the characteristics of campaigns (e.g., Chan and Parhankangas 2017; Parhankangas and Renko 2017; Skirnevskiy et al. 2017). Other work has focused on estimating causal effects using random assignment to experiments using differences in campaign characteristics as treatment effects (e.g., Allison et al. 2017; Stevenson et al. 2018; Younkin and Kuppuswamy 2018). Our study, in contrast, focuses on prediction to build models that control for confounding factors and explanatory variables to crowdfunding campaign success. With our approach, we extend existing research by new data and methods. The very novelty of this contribution is the volume of the dataset, comprising not only description but also speech and, in particular, video-related data at larger scale.

We derive dialectic particularities in text, speech, and video characteristics that determine whether or not campaigns are more likely to be successful. Detecting and understanding the influence that language and visual information have on the consumer’s perception of crowdfunding campaigns is difficult and complex, especially for the human observer. Our machine learning approach assists in detecting patterns that are difficult for humans to find, but the intuition derived from the model still requires human input to make the results accessible and to interpret them against the background of theoretical induction subsequently. Accordingly, we suggest that linguistic expression in text and speech that are abstract and more emotionally salient work well in increasing campaign success. The way information are

conveyed and construed is an area that warrants further attention for both theory building and empirical inferences.

As machine learning allows for “*algorithmic induction*,” it “*yields identical (or highly similar) conclusions when applied by different observers to the same data*” (Puranam et al. 2018, 1). Consequently, we believe our findings not to be sample-specific but rather generalizable across datasets. As such, our insights provided are both reproducible and robust to alternative variants of crowdfunding datasets used. In summary, we believe that the application of machine learning to entrepreneurship research brings about unprecedented opportunities and helps to tackle empirical and theoretical challenges that hitherto remain inconclusive for various reasons

9 Ethical considerations

For the training of the neural network, only *publicly* accessible information was analyzed, just as every human observer could perceive it. In order to respect personal rights, this study does not analyze or show any data that allows for conclusions about individual persons.

Acknowledgments We wish to thank Kexin Li, Hongzhu Chen, and Julius Scheuber for their help during the data aggregation process. The authors also acknowledge the RWTH Aachen (Germany) project house “ICT Foundations of a Digitized Industry, Economy, and Society” for supporting this research. C.H. gratefully acknowledges support by the Dr. Werner Jackstädt Stiftung. We further like to thank the German Federal Ministry of Education and Research for supporting the project within the framework of the exploratory research project “InnoFinance” (01IO1702) and Google Inc. for providing free access to the Cloud Video Intelligence API.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines*. Boston: Harvard Business Review Press.
- Ahlers, G. K., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice*, 39(4), 955–980.
- Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500.
- Allison, T. H., Davis, B. C., Short, J. C., & Webb, J. W. (2015). Crowdfunding in a prosocial microlending environment: Examining the role of intrinsic versus extrinsic cues. *Entrepreneurship Theory and Practice*, 39(1), 53–73.
- Allison, T. H., Davis, B. C., Webb, J. W., & Short, J. C. (2017). Persuasion in crowdfunding: An elaboration likelihood model of crowdfunding performance. *Journal of Business Venturing*, 32(6), 707–725.
- An, J., Quercia, D., & Crowcroft, J. (2014). Recommending investors for crowdfunding projects. In *Proceedings of the 23rd international conference on world wide web* (pp. 261–270). ACM.
- Anglin, A. H., Short, J. C., Drover, W., Stevenson, R. M., McKenny, A. F., & Allison, T. H. (2018). The power of positivity? The influence of positive psychological capital language on crowdfunding performance. *Journal of Business Venturing*, 33(4), 470–492.
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2), 3–32.
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489), 1390–1396.
- Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2), 29–50.
- Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). A neural probabilistic language model. *Journal of Machine Learning Research*, 3(2), 1137–1155.
- Bhattacharjee, A., & Sanford, C. (2006). Influence processes for information technology acceptance: An elaboration likelihood model. *MIS Quarterly*, 30(4), 805–825.
- Bird, S. (2006). NLTk: the natural language toolkit. In *Proceedings of the COL-ING/ACL on Interactive Presentation Sessions* (pp. 69–72). Association for Computational Linguistics.
- Biswas, D., & Biswas, A. (2004). The diagnostic role of signals in the context of perceived risks in online shopping: Do signals matter more on the web? *Journal of Interactive Marketing*, 18(3), 30–45.
- Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review*, 91(5), 63–72.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3(1), 993–1022.
- Brush, C. G., Manolova, T. S., & Edelman, L. F. (2008). Properties of emerging organizations: An empirical test. *Journal of Business Venturing*, 23(5), 547–566.
- Castano, R., Sujana, M., Kacker, M., & Sujana, H. (2008). Managing consumer uncertainty in the adoption of new products: Temporal distance and mental simulation. *Journal of Marketing Research*, 45(3), 320–336.
- Chafkin, M. (2015). Legendary angel investor Ron Conway isn’t looking at your idea, he’s looking at you. <https://www.>

- fastcompany.com/3043196/ron-conwayand-y-combinator-a-love-story. Accessed 12 Dec 2018.
- Chan, R., & Parhankangas, A. (2017). Crowdfunding innovative ideas: How incremental and radical innovativeness influence funding outcomes. *Entrepreneurship Theory and Practice*, *41*(2), 237–263.
- Chan, R., & Park, H. D. (2015). How images and color in business plans influence venture investment screening decisions. *Journal of Business Venturing*, *30*(5), 732–748.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery.
- Cordova, A., Dolci, J., & Gianfrate, G. (2015). The determinants of crowdfunding success: Evidence from technology projects. *Procedia-Social and Behavioral Sciences*, *181*, 115–124.
- Cumming, D., Leboeuf, G., Schwienbacher, A. (2015) Crowdfunding models: Keep-it-all vs. all-or-nothing. Available at SSRN. <https://doi.org/10.2139/ssrn.2447567>. Accessed 12 Dec 2018.
- Dahl, G.E., Adams, R.P., Larochelle, H. (2012) Training restricted Boltzmann machines on word observations. *arXiv*. <https://arxiv.org/abs/1202.5695>. Accessed December 12, 2019.
- Davis, B. C., Hmieleski, K. M., Webb, J. W., & Coombs, J. E. (2017). Funders' positive affective reactions to entrepreneurs' crowdfunding pitches: The influence of perceived product creativity and entrepreneurial passion. *Journal of Business Venturing*, *32*(1), 90–106.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009* (pp. 248–255). Institute of Electrical and Electronics Engineers.
- Desai, N., Gupta, R., Truong, K. (2015) Plead or pitch? The role of language in Kickstarter project success. *Report*. http://cs229.stanford.edu/proj2015/239_report.pdf. Accessed 12 Dec 2018.
- Dey, S., Duff, B. R., Karahalios, K., & Fu, W. T. (2017). The art and science of persuasion: Not all crowdfunding campaign videos are the same. In *The 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (pp. 755–769). Association for Computing Machinery.
- Dhar, R., & Kim, E.Y. (2007). Seeing the forest or the trees: Implications of construal level theory for consumer choice. *Journal of Consumer Psychology*, *17*(2), 96–100.
- Doyle, V., Freeman, O., O'Rourke, B.K. (2017) A multimodal discourse analysis exploration of a crowdfunding entrepreneurial pitch. In: *European Conference on Research Methodology for Business and Management Studies* (pp. 415–423). Academic Conferences International Limited.
- Du, Q., Fan, W., Qiao, Z., Wang, G., Zhang, X., Zhou, M. (2015) Money talks: A predictive model on crowdfunding success using project description. In: *Proceedings of Americas Conference on Information Systems*. <http://aisel.aisnet.org/amcis2015/BizAnalytics/GeneralPresentations/37>. Accessed December 12, 2019.
- Ernst, H., Hoyer, W. D., & Ru'bsaamen, C. (2010). Sales, marketing, and research-and- development cooperation across new product development stages: Implications for success. *Journal of Marketing*, *74*(5), 80–92.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115.
- Etter, V., Grossglauser, M., & Thiran, P. (2013). Launch hard or go home!: Predicting the success of Kickstarter campaigns. In *Proceedings of the First ACM Conference on Online Social Networks* (pp. 177–182). Association for Computing Machinery.
- Firth J. R. (1957). A synopsis of linguistic theory 1930-55. *Studies in Linguistic Analysis*. Reprinted in: Palmer, F. R. (ed.) (1968). *Selected Papers of J. R. Firth 1952-59* (pp. 168–205). Longmans, London.
- Fisher, G., Kotha, S., & Lahiri, A. (2016). Changing with the times: An integrated view of identity, legitimacy, and new venture life cycles. *Academy of Management Review*, *41*(3), 383–409.
- Gafni, H., Marom, D., & Sade, O. (2019). Are the life and death of an early-stage venture indeed in the power of the tongue? Lessons from online crowdfunding pitches. *Strategic Entrepreneurship Journal*, *13*(1), 3–23.
- George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, *57*(2), 321–326.
- Gerber, E. M., Hui, J. S., & Kuo, P. Y. (2012). Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms. In *Proceedings of the International Workshop on Design, Influence, and Social Technologies: Techniques, Impacts and Ethics* (Vol. 2, p. 11).
- Goode, M. R., Dahl, D. W., & Moreau, C. P. (2010). The effect of experiential analogies on consumer perceptions and attitudes. *Journal of Marketing Research*, *47*(2), 274–286.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning, vol 1*. Cambridge: MIT Press.
- Greenberg, M. D., Pardo, B., Hariharan, K., & Gerber, E. (2013). Crowdfunding support tools: predicting success & failure. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 1815–1820). Association for Computing Machinery.
- Hayes-Roth, B. F., Waterman, D. A. and Lenat, D. (1983). *Building Expert Systems*, Reading MA: Addison-Wesley, 1983.
- He K., Zhang X., Ren S. & Sun J. (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). Institute of Electrical and Electronics Engineers.
- von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science*, *32*(7), 791–805.
- Honig, B., & Karlsson, T. (2004). Institutional forces and the written business plan. *Journal of Management*, *30*(1), 29–48.
- Hui, J., Greenberg, M., & Gerber, E. (2013). Understanding crowdfunding work: Implications for support tools. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 889–894). Association for Computing Machinery.
- Jurafsky, D., Martin, J. H. (2016) *Speech and Language Processing*. Stanford University, <https://web.stanford.edu/~jurafsky/slp3/>. Accessed 12 Dec 2018.

- Kaminski, J., Jiang, Y., Piller, F., & Hopp, C. (2017). Do user entrepreneurs speak different? Applying natural language processing to crowdfunding videos. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2683–2689). Association for Computing Machinery.
- Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2009). Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *The Journal of Finance*, 64(1), 75–115.
- Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, 28(3), 25–48.
- Kim, P., Longest, K. C., & Lippmann, S. (2015). The tortoise versus the hare: Progress and business viability differences between conventional and leisure-based founders. *Journal of Business Venturing*, 30(2), 185–204.
- Kim J., Cho D., Lee B. (2016) The mind behind crowdfunding: An empirical study of speech emotion in fundraising success. In: Proceedings of the International Conference on Information Systems 2016. <https://aisel.aisnet.org/icis2016/Crowdsourcing/Presentations/23/>.
- Kirsch, D., Goldfarb, B., & Gera, A. (2009). Form or substance: The role of business plans in venture capital decision making. *Strategic Management Journal*, 30(5), 487–515.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). Human decisions and machine predictions. *The Quarterly Journal of Economics*, 133(1), 237–293.
- Koch, J. A., Cheng, Q. (2016) The role of qualitative success factors in the analysis of crowdfunding success: Evidence from kickstarter. Available at SSRN. <https://ssrn.com/abstract=2808428>. Accessed 12 Dec 2018.
- Lai C. Y., Lo P. C., Hwang S. Y. (2017) Incorporating comment text into success prediction of crowdfunding campaigns. In: *Proceedings of the Pacific Asia Conference on Information Systems 2017*, <https://aisel.aisnet.org/pacis2017/156>. Accessed 12 Dec 2018.
- Lau, J. H., Baldwin, T. (2016) An empirical evaluation of doc2vec with practical insights into document embedding generation. *arXiv*. <https://arxiv.org/abs/1607.05368>. Accessed December 12, 2019.
- Le Q., Mikolov, T. (2014) Distributed representations of sentences and documents. In: *Proceedings of the 31st International Conference on Machine Learning* (pp 1188–1196).
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436.
- Lee, S., Lee, K., & Hc, K. (2018). Content-based success prediction of crowdfunding campaigns: A deep learning approach. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 193–196). Association for Computing Machinery.
- Li, Y., Rakesh, V., & Reddy, C. K. (2016). Project success prediction in crowdfunding environments. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining* (pp. 247–256). Association for Computing Machinery.
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75(1), 5.
- Licklider, J. C. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, (1), 4–11.
- Lu, C. T., Xie, S., Kong, X., & Yu, P. S. (2014). Inferring the impacts of social media on crowdfunding. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining* (pp. 573–582). Association for Computing Machinery.
- Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2579–2605.
- Malone, T. W. (2018). *Superminds: The surprising power of people and computers thinking together*. New York: Little, Brown.
- Mikolov, T., Chen, K., Corrado, G., Dean, J. (2013a) Efficient estimation of word representations in vector space. *arXiv*. <https://arxiv.org/abs/1301.3781>. Accessed December 12, 2019.
- Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. (2013b) Distributed representations of words and phrases and their compositionality. In: *Advances in Neural Information Processing Systems* (pp 3111–3119).
- Mitra, T., Gilbert, E. (2014) The language that gets people to give: Phrases that predict success on Kickstarter. In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp 49–61). Association for Computing Machinery.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529.
- Mollick, E. (2013) Swept away by the crowd? Crowdfunding, venture capital, and the selection of entrepreneurs. Available at SSRN. <https://doi.org/10.2139/ssrn.2239204>. Accessed December 12, 2019.
- Mollick E. & Kuppuswamy, V. (2014a) After the campaign: Outcomes of crowdfunding. Available at SSRN, <http://doi.org/10.2139/ssrn.2376997>. Accessed December 12, 2019.
- Mollick, E. (2014b). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, 29(1), 1–16.
- Mollick, E. (2015) Delivery rates on Kickstarter. Available at SSRN. <https://doi.org/10.2139/ssrn.2699251>. Accessed December 12, 2019.
- Mollick, E., & Nanda, R. (2015). Wisdom or madness? Comparing crowds with expert evaluation in funding the arts. *Management Science*, 62(6), 1533–1553.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106.
- Nair V., Hinton G. E. (2010) Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning* (pp 807–814).
- Oo, P. P., Allison, T. H., Sahaym, A., & Juasrikul, S. (2018). User entrepreneurs' multiple identities and crowdfunding performance: Effects through product innovativeness, perceived passion, and need similarity. *Journal of Business Venturing* Forthcoming.
- Parhankangas, A., & Ehrlich, M. (2014). How entrepreneurs seduce business angels: An impression management approach. *Journal of Business Venturing*, 29(4), 543–564.

- Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing*, 32(2), 215–236.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(10), 2825–2830.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001* (Vol. 71(2001), p. 2001). Mahwah: Lawrence Erlbaum Associates.
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.
- Puranam, P., Shrestha, Y. R., He, V. F., von Krogh, G. (2018) Algorithmic induction through machine learning: Opportunities for management and organization research. Available at SSRN. <https://doi.org/10.2139/ssrn.3140617>. Accessed 12 Dec 2018.
- Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton: Princeton University Press.
- Sako, M. (1992). *Price, quality and trust: Inter-firm relations in Britain and Japan* (Vol. 18). Cambridge: Cambridge University Press.
- Scheaf, D. J., Davis, B. C., Webb, J. W., Coombs, J. E., Borns, J., Holloway, G., et al. (2018). Signals' flexibility and interaction with visual cues: Insights from crowdfunding. *Journal of Business Venturing*, 33(6), 720–741.
- Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., et al. (2017) 2017 AI index report. <http://cdn.aiindex.org/2017-report.pdf>. Accessed 12 Dec 2018.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587), 484–489.
- Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., et al. (2017). Mastering the game of go without human knowledge. *Nature*, 550(7676), 354.
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419), 1140–1144.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.
- Skirnevskiy, V., Bendig, D., & Brettel, M. (2017). The influence of internal social capital on serial creators' success in crowdfunding. *Entrepreneurship Theory and Practice*, 41(2), 209–236.
- Socher R., Lin C. C., Manning C., Ng A.Y. (2011) Parsing natural scenes and natural language with recursive neural networks. In: *Proceedings of the 28th International Conference on Machine Learning* (pp 129–136).
- Řehůřek, R., Sojka, P. (2010) Software framework for topic modelling with large corpora. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (pp 45–50). <http://is.muni.cz/publication/884893/en>. Accessed 12 Dec 2018.
- Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1), 11–21.
- Stanko, M. A., & Henard, D. H. (2017). Toward a better understanding of crowdfunding, openness and the consequences for innovation. *Research Policy*, 46(4), 784–798.
- Stevenson, R. M., Ciuchta, M. P., Letwin, C., Dinger, J. M., & Vancouver, J. B. (2018). Out of control or right on the money? Funder self-efficacy and crowd bias in equity crowdfunding. *Journal of Business Venturing*.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24–54.
- Tran T, Dontham MR, Chung J, Lee K (2016) How to succeed in crowdfunding: A long-term study in Kickstarter. *arXiv*. <https://arxiv.org/abs/1607.06839>. Accessed December 12, 2019.
- Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, 110(3), 403.
- Varian, H. R. (2014). Big data: New tricks for econometrics. *The Journal of Economic Perspectives*, 28(2), 3–27.
- von Krogh, G. (2018). Artificial intelligence in organizations: New opportunities for phenomenon-based. Theorizing. *Academy of Management Discoveries* 4(4), 404–409. <https://doi.org/10.5465/amd.2018.0084> Accessed 22 Dec 2018.
- Wang, S., & Manning, C. D. (2012). Baselines and bigrams: Simple, good sentiment and topic classification. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics* (pp. 90–94). Association for Computational Linguistics.
- Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., et al. (2016) Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv*. <https://arxiv.org/abs/1609.08144>. Accessed December 12, 2019.
- Younkin P. & Kuppuswamy V. (2018) Discounted: The effect of founder race on the price of new products. *Journal of Business Venturing*, 34(2), 389–412.
- Yuan, H., Lau, R. Y., & Xu, W. (2016). The determinants of crowdfunding success: A semantic text analytics approach. *Decision Support Systems*, 91, 67–76.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.