Let's Make Your Request More Persuasive: Modeling Persuasive Strategies via Semi-Supervised Neural Nets on Crowdfunding Platforms

Diyi Yang*, **Jiaao Chen***, **Zichao Yang**, **Dan Jurafsky**, **Eduard Hovy**
Georgia Institute of Technology, Carnegie Mellon University, Stanford University
diyi.yang@cc.gatech.edu
{jiaaoc, zichaoy, hovy}@andrew.cmu.edu
jurafsky@stanford.edu

Abstract

Modeling what makes a request persuasive—eliciting the desired response from a reader—is critical to the study of propaganda, behavioral economics, and advertising. Yet current models can't quantify the persuasiveness of requests or extract successful persuasive strategies. Building on theories of persuasion, we propose a neural network to quantify persuasiveness and identify the persuasive strategies in advocacy requests. Our semi-supervised hierarchical neural network model is supervised by the number of people persuaded to take actions and partially supervised at the sentence level with human-labeled rhetorical strategies. Our method outperforms several baselines, uncovers persuasive strategies—offering increased interpretability of persuasive speech—and has applications for other situations with document-level supervision but only partial sentence supervision.

1 Introduction

Crowdfunding platforms are a popular way to raise funds for projects. For example, Kiva, a peer-to-peer lending platform, has crowd-funded more than a million loans, totaling over $1 billion since 2005. Kickstarter, another online crowdfunding platform, successfully funded 110,270 projects with a total of over 2 billion dollars. Yet most projects still suffer from low success rates. How can we help requesters craft persuasive and successful pitches to convince others to take actions?

Persuasive communication has the potential to shape and change people's attitudes and behaviors (Hovland et al., 1953), and has been widely researched in various fields such as social psychology, marketing, behavioral economics, and political campaigning (Shrum et al., 2012). One of the most influential theories in the advertising literature is Chaiken's systematic-heuristic dual processing theory, which suggests that people process persuasive communication by evaluating the quality of arguments or by relying on inferential rules. Some such heuristic rules are commonly used in consumer behaviors; commercial websites may highlight the limited availability of their items "*In high demand - only 2 left on our site!*" or emphasize the person in authority "*Speak to our head of sales—he has over 15 years' experience selling properties*" to attract potential consumers. Although numerous studies on persuasion have been conducted (Chaiken, 1980), we still know little about the way how persuasion functions in the wild and how it can be modeled computationally.

In this work, we utilize neural-network based methods to computationally model persuasion in requests from crowdfunding websites. We build on theoretical models of persuasion to operationalize persuasive strategies and ensure generalizability. We propose to identify the persuasive strategy employed in each sentence in each request. However, constructing a large dataset with persuasion strategies labeled at the sentence level is time-consuming and expensive. Instead, we propose to use a small amount of hand-labeled sentences together with a large number of requests automatically labeled at the document level by the number of persuaded support actions. Our model is a semi-supervised hierarchical neural network that identifies the persuasive strategies employed in each sentence, where the supervision comes from the overall persuasiveness of the request. We propose that the success of requests could have substantive explanatory power to uncover their persuasive strategies. We also introduce an annotated corpus with sentence-level persuasion strategy labels and document-level persuasiveness labels, to facilitate future work on persuasion. Experiments

*Equal contribution. This work was done when the first two authors were students at CMU.

show that our semi-supervised model outperforms several baselines. We then apply this automated model to unseen requests from different domains and obtain nuanced findings of the importance of different strategies on persuasion success. Our model can be useful in any situation in which we have exogenous document-level supervision, but only small amounts of expensive human-annotated sentence labels.

2 Related Work

Computational argumentation has received much recent attention (Ghosh et al., 2016; Stab and Gurevych, 2017; Peldszus and Stede, 2013; Stab et al., 2018; Ghosh et al., 2014). Most work has either identified the arguments in news articles (Sardianos et al., 2015) or user-generated web content (Habernal and Gurevych, 2017; Musi et al., 2018), or classified argument components (Zhang and Litman, 2015) into claims and premises, supporting and opposing claims, or backings, rebuttals and refutations . For example, Stab and Gurevych (2014) proposed structural, lexical, syntactic and contextual features to identify convincing components of Web arguments including claim, major claim, and premise. Similarly, Zhang and Litman (2015) studied student essay revisions and classified a set of argumentative actions associated with successful writing such as warrant/reasoning/backing, rebuttal/reservation, and claims/ideas. Habernal and Gurevych (2016) investigated the persuasiveness of arguments in any given argument pair using bidirectional LSTM. Hidey et al., (2017) utilized the persuasive modes—ethos, logos, pathos—to model premises and the semantic types of argument components in an online persuasive forum.

While most computational argumentation focuses on the relational support structures and factual evidence to make claims, **persuasion** focuses more on language cues aimed at shaping, reinforcing and changing people's opinions and beliefs. How language changes people's attitudes and behaviors have received less attention from the computational community than argumentation, although there have been important preliminary work (Persing and Ng, 2017; Carlile et al., 2018). Farra et al., (2015) built regression models to predict essay scores based on features extracted from opinion expressions and topical elements. Chatterjee et al., (2014) used verbal descriptors and para-verbal markers of hesitation to predict speakers' persuasiveness on website housing videos of product reviews. When looking at persuasion in the context of online forum discussions (Wei et al., 2016), Tan et al., (2016) found that on the Change My View subreddit, interaction dynamics such as the language interplay between opinion holders and other participants provides highly predictive cues for persuasiveness. Using the same dataset, Wel et al., (2016) extracted a set of textual information and social interaction features to identify persuasive posts.

Recently, Pryzant et al., (2017) introduced a neural network with an adversarial objective to select text features that are predictive of some outcomes but decorrelated with others and further analyzed the narratives highlighted by such text features. Further work extended the model to induce narrative persuasion lexicons predictive of enrollment from course descriptions and sales from product descriptions (Pryzant et al., 2018a), and the efficacy of search advertisements (Pryzant et al., 2018b). Similar to their settings, we use the outcomes of a persuasive description to supervise the learning of persuasion tactics, and our model can similarly induce lexicons associated with successful narrative persuasion by examining highly attentional words associated with persuasion outcomes. Our work differs both in our semi-supervised method and also because we explicitly draw on the theoretical literature to model the persuasion strategy for each sentence in requests, allowing requests to have multiple persuasion strategies; our induced lexicons can thus be very specific to different persuasion strategies.

Other lines of persuasion work predict the success of requests on peer-to-peer lending or crowd-funding platforms, and mainly exploit request attributes like project description (Greenberg et al., 2013), project videos (Dey et al., 2017), and social predictors such as the number of backers (Etter et al., 2013) or specific types of project updates (Xu et al., 2014). Among them, only a few investigated the effect of language on the success of requests. Althoff et al., (2014) studied donations in Random Acts of Pizza on Reddit, using the social relations between recipient and donor plus linguistic factors to predict the success of these altruistic requests. Based on a corpus of 45K crowd-funded projects, Mitra and Gilbert (2014) found that 9M phrases commonly present in crowd-funding have reasonable predictive power in accounting for variance around successful funding, suggesting that language does exhibit some general principles of persuasion. Although this prior work offers predictive and insightful models, most studies chose their persuasion labels or variables without reference to a taxonomy of persuasion techniques nor to a principled method of choosing them. Some exceptions include Yang and Kraut (2017), Dey et al., (2017), and Rosenthal and McKeown (2017). For example, Yang and Kraut (2017) looked at the effectiveness of a set of persuasive cues in Kiva requests and found that certain heuristic cues are positively correlated with lenders' contribution.

Inspired by these prior work, we operationalize persuasive strategies based on theories of persuasion and aim to learn local structures/labels of sentences based on the global labels of paragraphs/requests. Our task is different from most previous work on semi-supervised learning for NLP (Liang, 2005; Yang et al., 2017) that focuses on the setting with partial data labels. While in computer vision, there is a lot of prior work in using image global labels to uncover local pixel level labels and bounding boxes of objects (Oquab et al., 2015; Pinheiro and Collobert, 2015), the investigation of this task in NLP, to the best of our knowledge, is novel and could potentially have much broader applications.

3 Research Context

We situate this research within the team forums of Kiva[1], the largest peer-to-peer lending website. These self-organized lending teams are built around common interests, school affiliation or location. In such teams, members can post messages in their team discussion board to persuade other members to lend to a particular borrower. One such message is shown in Figure 1. A borrower, Sheila, posted a message on Kiva to request loans for woman-led group. As highlighted in the figure, she made use of several persuasion strategies such as commitment, concreteness, and impact to render her request more persuasive. We define the *persuasiveness* score of a request message as the number of team members (in log-scale) who read the message and make loans to the mentioned borrower. We then regard this overall persuasiveness of messages as high-level supervision for training our model to determine which persuasion strategy is used in each sentence inside each message.

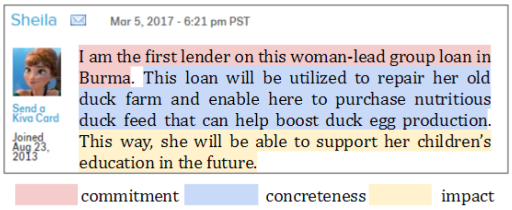


Figure 1: An anonymized advocating message that persuaded 5 members to lend to the mentioned borrower. Persuasion strategies are highlighted.

4 Persuasion Strategies

Numerous studies have investigated the basic principles that govern getting compliance from people (Cialdini and Garde, 1987; Petty et al., 1983). In this work, we utilized Chiaken's 1980 systemic-heuristic model of social information processing, which suggests that people process persuasive requests by assessing the quality of arguments (systematic processing) or by relying on heuristic rules (heuristic processing). Building on that, we first borrow several commonly used heuristic principles (Cialdini and Garde, 1987) that are also suitable for our context as below.

- **Scarcity** states that people tend to value an item more as soon as it becomes rare, distinct or limited. For example, take the use of 'expire'in this message: *"This loan is going to expire in 35 mins..."*.
- The principle of **Emotion** says that making messages full of emotional valence and arousal affect (e.g., describing a miserable situation or a happy moment) can make people care and act, e.g., "*The picture of widow Bunisia holding one of her children in front of her meager home brings tears to my eyes..*", similar to *Sentiment* and *Politeness* used by Althoff et al., (2014) and Tan et al., (2016), and *Pathos* used by Hidey et al., (2017).
- **Commitment** states that once we make a choice, we will encounter pressures that cause us to respond in ways that justify our earlier decision, and to convince others that we have made the correct choice. Here it could be mentioning their contribution in the message, e.g., *"I loaned to her already."*
- **Social Identity** refers to people's self-concept of their membership in a social

[1] https://www.kiva.org/

group, and people have an affinity for their groups over others, similar to *name mentions* in Rosenthal and McKeown (2017). Thus if a loan request comes from their own groups, they are more likely to contribute, such as *"For those of you in our team who love bread, here is a loan about bakery."*

- **Concreteness** refers to providing concrete facts or evidence, such as "*She wishes to have a septic tank and toilet, and is 51% raised and needs $825*", similar to *Claim* and *Evidence* (Zhang et al., 2016; Stab and Gurevych, 2014)), *Evidentiality* (Althoff et al., 2014), and *Logos* (Hidey et al., 2017).

We also propose a new strategy to capture importance or impact on these requests:

- **Impact and Value** emphasizes the importance or bigger impact of this loan, such as *"... to grow organic rice. Then, she can provide better education for her daughter"*.

Note that other persuasion tactics such as *Reciprocity* — "feel obligated to return something after receiving something of value from another" — and *Authority* — "comply with the requests of authority in an unthinking way to guide their decisions"— are also widely used in persuasive communication. However, in this context, we did not observe enough instances of them.

5 Semi-supervised Neural Net

Given a message $M = \{S_0, S_1, ..., S_L\}$ consisting of L sentences that the author posted to advocate for a loan, our task is to predict the persuasion strategies p_i employed in each sentence S_i, $i \in [0, L]$. However, purely constructing a large-scale dataset that contains such labels of sentence-level persuasion strategy is often time-consuming and expensive. Instead, we propose to utilize a small amount of labeled and a large amount of unlabeled data. We design a semi-supervised hierarchical neural network to identify the persuasive strategies employed in each sentence, where the supervision comes from the sentence-level labels g in a small portion of data and the overall persuasiveness scores y of messages. The overall architecture of our method is shown in Figure 2.

5.1 Sentence Encoder

Given a sentence S_i with words $w_{i,j}$, $j \in [0, l]$ and l is the sentence length, a GRU (Bahdanau et al., 2014) is used to incorporate contextual cues of words into hidden state $h_{i,j}$. This GRU reads the sentence S_i from $w_{i,1}$ to $w_{i,l}$ and encodes each word $w_{i,j}$ with its context into hidden state $h_{i,j}$:

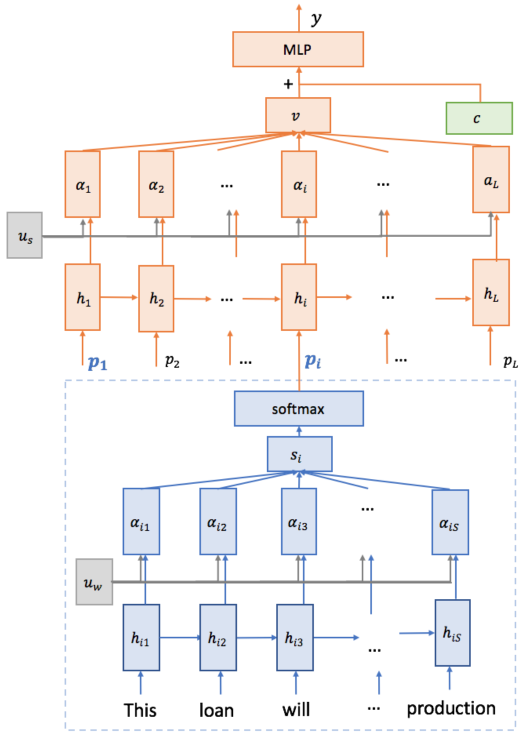


Figure 2: The overall model architecture. The blue part describes the sentence encoder. Sentences with labels of persuasion strategies are highlighted with dark blue like p_1. The orange part shows the document encoder.

$$h_{i,j} = \text{GRU}(W_e w_{i,j}, h_{i,j-1}), j \in [0, l]. \quad (1)$$

where W_e is the word embedding matrix. To learn the characteristic words associated with the persuasive strategy in a sentence, we apply an attention mechanism (Bahdanau et al., 2014; Yang et al., 2016). The representation of those words are then aggregated to form the sentence vector s_i. We formulated this word level attention as follows:

$$u_{i,j} = \tanh(W_w h_{i,j} + b_w) \quad (2)$$

$$\alpha_{i,j} = \frac{\exp(u_{i,j}^\intercal u_w)}{\sum_k \exp(u_{i,k}^\intercal u_w)} \quad (3)$$

$$s_i = \sum_j \alpha_{i,j} h_{i,j} \quad (4)$$

where u_w is a context vector that queries the characteristic words associated with different persuasion strategies. It is randomly initialized and jointly learned from data.

5.2 Latent Persuasive Strategies

We assume that each sentence instantiates only one type of persuasion strategy. For example, a sentence *"She is 51% raised and needs $825 in 3 days"* employs *Scarcity*, trying to emphasize limited time availability. We propose to use the high level representation of each sentence to predict the latent variable:

$$p_i = \text{softmax}(W_v s_i + b_v) \tag{5}$$

5.3 Document Encoder

After obtaining the sentence vector p_i, we can get a document vector in a similar way:

$$h_i = \text{GRU}(p_i, h_{i-1}), i \in [0, L] \tag{6}$$

where L denotes the number of sentences in a message. Similarly, we introduce an attention mechanism to measure the importance of each sentence and its persuasion strategy via a context vector u_s:

$$u_i = \tanh(W_s h_i + b_s) \tag{7}$$

$$\alpha_i = \frac{\exp(u_i^\intercal u_s)}{\sum_k \exp(u_k^\intercal u_s)} \tag{8}$$

$$v = \sum_i \alpha_i h_i \tag{9}$$

5.4 Semi-Supervised Learning Objective

The document vector v is a high-level representation of the document and can be used as a set of features for predicting $\tilde{y}$, the persuasiveness of a message, i.e., how many team members will make loans to the project mentioned in this message. We also include a context vector c to further assist the prediction of making loans. For instance, c could represent the number of team members in a team, the total amount of money contributed by this team in the past, etc.

$$\tilde{y} = W_f[v, c] + b_f \tag{10}$$

We then can use the mean squared error between the predicted and ground truth persuasiveness as training loss. To take advantage of the labeled subset that has sentence level annotation of persuasive strategies, we reformulate this problem as a semi-supervised learning task:

$$l = \gamma \sum_{d \in C_L} (y_d - \tilde{y_d})^2 - \beta \sum -g_i \log p_i \tag{11}$$

$$+ (1-\gamma) \sum_{d' \in C_U} (y_{d'} - \tilde{y_{d'}})^2 \tag{12}$$

Here, C_L refers to the document corpus with sentence level persuasion labels. C_U denotes those without any sentence labels. g_i refers to the persuasion strategy in sentence S_i, and p_i is predicted by our model. γ and β are used as re-weight factors to trade off the penalization and reward introduced by different components.

6 Experiment

6.1 Dataset

Our collaboration with Kiva provided us access to all public data dumps of the team discussion forums on Kiva. Here we only focused on messages that have explicit links because in most cases, members need to include the loan link to better direct others to a specific loan or borrower. After removing messages that do not contain any links, we obtained 41,666 messages that contain loan advocacy. We used Amazon's Mechanical Turk (MTurk) to construct a reliable, hand-coded dataset to obtain the persuasion strategy label for each sentence. To increase annotation quality, we required Turkers to have a United States location with 98% approval rate for their previous work on MTurk. Since messages often contain different numbers of sentences, which might be associated with different sets of persuasion strategies, we sampled 200 messages for each fixed message length ranging from one sentence to six sentences, in order to guarantee that our hand-coded dataset reasonably represents the data. Messages with at most six sentences accounted for 89% percentages among all messages in our corpus. Each sentence in a message was labeled by two Mechanical Turk Master Workers [2]. To assess the reliability of the judges' ratings, we computed the intra-class correlation (ICC), and obtained an overall ICC score of 0.524, indicating moderate agreement among annotators (Cicchetti, 1994). The distribution for each persuasion strategy in the annotated corpus is described in the blue line in Figure 3. We assigned a persuasion label to a sentence if two annotators gave consistent labels for it, and filtered out sentences that annotators disagreed on the label.

In the final annotated corpus, there were 1200 messages, with 2898 sentences. The average number of sentences is 2.4 and the average number of words per sentence is 17.3. For predicting the persuasive strategy in each sentence, we randomly

[2] https://www.mturk.com/worker/help:What-is-a-Mechanical-Turk-Master-Worker

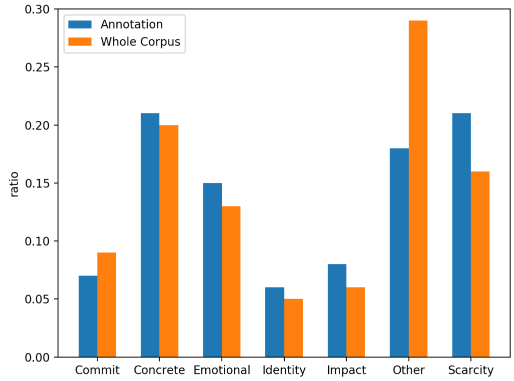


Figure 3: The distribution of each persuasion strategy in the annotation corpus and in the whole unlabeled corpus after prediction.

split 80% of this annotated corpus as the training set (2271 sentences in 1060 messages), 10% as the validation set (322 sentences in 70 messages), and 10% as the testing set (305 sentences in 70 messages). To further utilize supervision from the persuasiveness score of each message, we merged 1060 documents with sentence labels and 40,466 unlabeled messages, using it as the final training set for training semi-supervised models.

6.2 Model Setup and Baselines

We split documents into sentences and tokenize each sentence using Stanford's CoreNLP (Manning et al., 2014). Words appearing less than 5 times were replaced with a special UNK token. We trained the hyperparameters of the models on the validation set using Adam (Kingma and Ba, 2014). Specifically, we set the word embedding dimension to be 128, where the word embeddings are initialized randomly, and GRU dimension to be 256. The learning rate is set to be 5e-5. The balancer γ is the ratio of labeled data in a batch of training data. The balancer β is selected via grid search, searching in a set of (5, 10, 20, 50, 100), resulting in β=10.

We propose several baselines to predict the sentence level persuasion strategies for comparison with our model. (1) ***SVM + BoW*** is a *SVM* classifier with *RBF* kernel using bag-of-words features (one-hot). (2) ***GRU*** uses the hidden state at the last word as features to classify persuasive strategies, a special case of our *SH Net* model without the supervision from the overall persuasiveness scores. (3) ***bi-GRU*** uses bi-directional GRU.

H Net is a hierarchical GRU for classifying strategies with the supervision from the overall persuasiveness scores as shown in Figure 2, but it only adopts all the annotated messages. We denote our semi-supervised hierarchical model as **SH Net** (Semi Hierarchical Net), which utilizes both annotated messages and unlabeled corpus. **Semi-Att Net** builds on **SH Net** by incorporating both word-level and sentence-level attention. In addition to the textual cues in the advocation message, persuasive requests also depend on the context. We introduced a set of contextual descriptors into our semi-supervised hierarchical network, denoted as *SH-Att Plus Net*. Such features include the number of borrowers in this message, the number of team members in a team, the total amount of money contributed by this team, the number of messages ever posted in the discussion board of this team, and the amount of money requested in this loan.

6.3 Results

We evaluated the baselines and our hierarchical neural network models using accuracy, macro-averaged F1 score, macro-averaged precision and macro-averaged recall, as well as RMSE for evaluating the message level persuasiveness score prediction. As we can see in Table 1, when predicting the persuasive strategies (6 types of persuasive strategies plus an *Other* strategy), BoW + SVM gives a performance of 0.347 and a macro F1 of 0.229. A direct neural network *GRU* boosted the accuracy to 0.518, demonstrating the effectiveness of neural networks for sentence classification. When bi-directional contextual information is used, the sentence level prediction performance is 52.1%. Our hierarchical neural network achieved an accuracy of 48.2% and a macro F1 of 0.432. When incorporating the whole corpus of unlabeled messages, our semi-supervised neural network achieved an accuracy of 56.1% (16.4% improvement over *H Net*). This indicates that our semi-supervised model effectively takes advantage of the supervision from the small amount of labeled data and the overall persuasiveness scores. Moreover, we noticed that this semi-supervised neural network not only helps predict the sentence level persuasion strategies, but also assist the prediction of messages' overall persuasiveness with a 9% RMSE decrease. *Semi-Att* outperformed *SH Net* with an accuracy of 56.9%, and a macro F1 score of 0.518. Although the improvement from attention is minor (but significant), it's important for visualizing associations between words, persuasion strategies and persua-

	Evaluating Sentence Level Strategies				Doc Level
Model	**Accuracy**	**Macro F1**	**Macro Precision**	**Macro Recall**	**RMSE**
SVM (RBF) + BoW	0.347	0.229	0.364	0.167	-
GRU	0.518	0.479	0.479	0.479	-
bi-GRU	0.521	0.440	0.445	0.436	-
Hierarchical Net (H Net)	0.482	0.432	0.430	0.432	1.15
Semi Net (SH Net)*	0.561	0.513	0.504	0.522	1.05
Semi-Att Net*	**0.569**	**0.518**	**0.512**	**0.534**	1.04
Semi-Att Plus Net	0.552	0.513	0.515	0.512	**0.87**

Table 1: Results of different models. * indicates that the model is significantly better than the one above it.

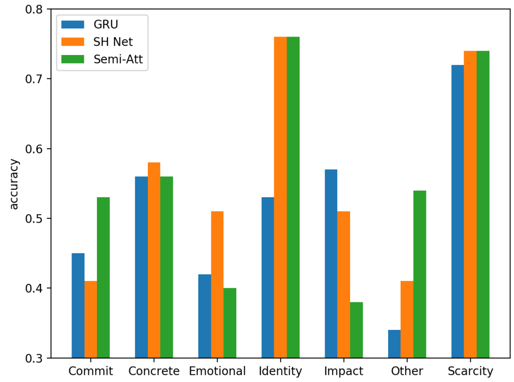


Figure 4: The accuracy for each persuasion strategy evaluated via *GRU*, *SH Net* and *Semi-Att*.

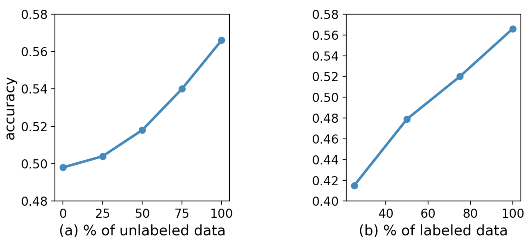


Figure 5: Model performances with different portion of unlabeled data (a) and labeled data (b).

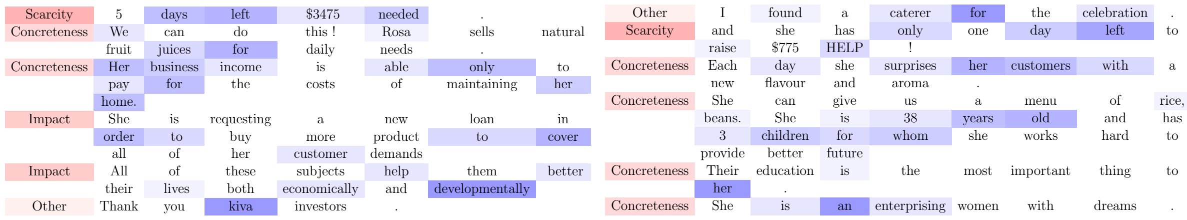
sion outcomes. Interestingly, incorporating contextual descriptors did not help the prediction of persuasion strategies. However, such contextual information strongly predicted the overall persuasiveness, decreasing RMSE to 0.84 from 1.04.

Strategy-Level Performance: We also report the accuracy per persuasion strategy category via *Semi-Att*, *SH Net* and simple *GRU* in Figure 4. It seems that overall neural models are better at capturing persuasion strategies such as concreteness, identity and scarcity. This might be because people are concrete by using specific terms such as numbers or entities that are easy to model. Similar principles might also occur for social identity and scarcity where the use of words such as "*we*", "*our*" and "*expire*", "*left*" can reveal a lot about the persuasion strategies.

Strategy	Top Ranked Keywords
Commitment	joined, lenders, loaning, lend, loan just, join, loaned, made, lent
Concreteness	women, married, old, heads, year-old money, sells, years, business, number
Emotion	hard, thank, better, grief, great maybe, help, please, thanks, happy
Identity	promotion, shall, captain, form, number spirits, lenders, member, team
Impact	improve, new, better, products, money to, use, business, more, order
Scarcity	minutes, there's, now, soon, go expire, hours, days, number, left

Table 2: Top ranked keywords for persuasion strategies

Different Percentage of Labeled Data: To figure out the importance of supervision from massages' overall persuasiveness scores, we experiment on *SH Net* with all the labeled messages. To this end, we include all the labeled messages, and vary the percentage of unlabeled corpus from 0%, 25%, 50%, 75%, to 100%, in Figure 5 (a). We found that as the amount of unlabeled messages increases, the accuracy of sentence level prediction increases as well, which further validates the effectiveness of the semi-supervised setting for persuasion strategy prediction. Similarly, to investigate the predictive power introduced by the sentence level labels, we also vary the percentage of labeled messages from 25%, 50%, 75%, to 100% when including the whole unlabeled corpus, as shown in Figure 5 (b). As expected, having more training data about sentence-level annotation increases the prediction performance. Overall, these experiments demonstrate the effectiveness of semi-supervised models for predicting sen-



(a) Predicted persuasiveness: 2.56 (after natural logarithm)

(b) Predicted persuasiveness: 1.75

Figure 6: Attention Visualization

tence level persuasion strategies. This enables us to obtain sentence level labels for any given paragraphs by using a small amount of labeled data.

6.4 Visualization

To validate whether our semi-supervised model captures characteristic words and sentences in requests, we visualize the attention in a sentence in Figure 6. We show the predicted persuasion label for each sentence in a message in red, with the color scale indicating its learned attention weight. Word-level attention is highlighted in blue. As we can see in Figure 6(b), our model places emphasis on *Scarcity*, and highlights words such as "*left*" and "*day*" that carry the scarcity meaning. Similarly, in the second message— *5 days left* 3475 *needed*—our model first labeled the sentence as *Scarcity*, and then picked words such as "*days*" and "*left*". Sentences that were predicted as *Concreteness* seem to contain specific entities and concepts such as "*business*", "*her*", and "*home*". For *Impact*, our model accurately localizes the words "*in order to*", "*cover*", and "*developmentally*".

To demonstrate that our model can learn representative words associated with different persuasion strategies, we show the 10 highest-scoring words from sentences with different labels in Table 2. Interestingly, *Commitment* is highly associated with words such as "*made*" and "*loan*". Explicit mentions of "*thanks*" and "*hard*" were found in sentences with *Emotional* labels. Sentences that emphasize their "*team*" as a whole were labeled as *Identity*. Overall, this validates that our model is able to select informative words associated with different persuasion strategies.

For further illustration, we visualized the attention weight distributions of different persuasion strategies. Since the number of sentences inside each message is intertwined with attention weights, we only plotted the distributions for messages with two or three sentences in Figure 7. We

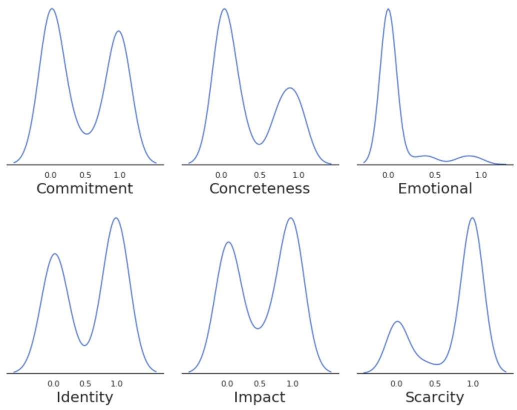


Figure 7: Attention weight distributions of persuasion strategies in requests with 2-3 sentences.

observed that *Scarcity*, *Identity*, and *Impact* seem to play a relatively more important role for influencing the success of requests, whereas *Emotional* language, *Commitment* and *Concreteness* seem to concentrate more on the lower weight ranges.

7 Importance of Persuasive Strategies

After applying the semi-supervised hierarchical neural network to the unlabeled 40552 messages, we obtained their sentence-level persuasive strategies usages. We showed the distribution of each persuasive strategy in the whole corpus in Figure 3, as described by the orange line. To further investigate how important each persuasive strategy is for convincing others to make loans, in this section, we present results on which of them are predictive via linear regression. All variables are standardized before entering the regression model. We controlled for the number of team members in a team, the total amount of money contributed by this team, the number of messages posted in the discussion board of this team. Since those variables are highly correlated with each other, we averaged them into a single variable to capture these team level attributes. We also controlled for the

Persuasion Strategy	**Kiva** (Coef.)	**RAOP** (Odds ratio)
Concreteness	0.041***	1.111***
Commitment	-0.015**	1.062
Emotional	0.030***	1.145***
Identity	0.087***	1.104**
Impact/Value	0.024***	1.084*
Scarcity	-0.076***	1.118***

Table 3: The influences of different persuasive strategies on request success on Kiva and RAOP. Here, p<0.001:***; p<0.01:**; p<0.05:*.

amount of money the borrower requested. We represented each message as a 6-dimensional vector to capture the amount of each persuasive strategy, which is calculated by selecting the maximum probability associated with each strategy from all sentences in this message. The persuasive strategy features significantly improve the model fit, as indicated by a 11.8% improvement in adjusted R-squared from 0.152 to 0.170. To demonstrate the generalizability of our persuasion strategies and the resulted semi-supervised model, we also applied our *Semi-Att* model to 5671 textual requests for pizza from the Reddit community "Random Acts of Pizza" (RAOP). Specifically, we used the data released by Althoff et al., (2014) where each request asked for a free pizza and the outcome whether its author received a pizza or not was provided in the dataset. Via *Semi-Att*, we were able to obtain the persuasive strategy used in each sentence of each request. Similarly, we built a logistic regression model to predict whether a request will receive the pizza or not, controlling for the community age of the requester, the number of subreddits the requester participated in, his/her number of posts as well as the votes (upvotes - downvotes) this requester had received.

As shown in the column of *RAOP* in Table 3, concreteness is significantly correlated with success on both datasets. This demonstrated that providing more evidence might help readers know the situation better, consistent with the effect of *Evidentiality* in Althoff et al., (2014). Similarly, making the request full of emotions (β=0.030, Odds ratio (OR) =1.145), mentioning the similarity between potential readers and the requester (β=0.087, *OR*=1.104), and talking about the potential impact and value for others (β=0.024, *OR*=1.084) are all significantly associated with increases in the persuasiveness of these requests across two contexts. In contrast, highlighting the urgency of the requests and emphasizing existing contribution to loans (β=-0.015) negatively correlate with request success (β=-0.076) on Kiva, confirming prior work (Yang and Kraut, 2017). This communicates to us that some of those loans might have expired before others read the request and took action given the limited time available, or it could be that members thought their actions might not help if the remaining money needed is high and the time left is low, different from the "limited-time offer" tactics widely used in commercial advertising. To sum up, the two analyses demonstrated that certain persuasive strategies such as *Identity* and *Impact* are consistently effective across two datasets, whereas *Scarcity* and *Commitment* contribute differently and need to be used with caution for different contexts.

8 Conclusion

In this work, we operationalized a set of persuasive strategies widely used in micro-lending platforms based on theories of persuasion, and developed an annotated corpus for identifying persuasion strategies. We designed a semi-supervised hierarchical neural network to identify the persuasive strategies contained in loan requests. Results show that our model improves accuracy considerably. We also showed how different persuasive strategies contribute to request success. In the future, we plan to build a richer taxonomy of persuasion strategies and incorporate additional neural architectures such as variational autoencoders to better represent sentences in each message to further assist the modeling of persuasiveness. Beyond the text, images and even audios may provide additional insights on the successes of persuasive requests. A more generalized persuasion framework is also needed to jointly learn persuasion strategies in different domains. Our model also has important applications to other domains, such as in computational advertisements, micro-funding platforms and political campaigns.

Acknowledgement

The authors would like to thank Jason Eisner for his help at the brainstorm stage and insightful followup suggestions, and the anonymous reviewers for their helpful comments. Diyi Yang was supported by Facebook Fellowship.

References

Tim Althoff, Cristian Danescu-Niculescu-Mizil, and Dan Jurafsky. 2014. How to ask for a favor: A case study on the success of altruistic requests. In *Proceedings of ICWSM*.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Winston Carlile, Nishant Gurrapadi, Zixuan Ke, and Vincent Ng. 2018. Give me more feedback: Annotating argument persuasiveness and related attributes in student essays. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 621–631.

Shelly Chaiken. 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology*, 39(5):752.

Moitreya Chatterjee, Sunghyun Park, Han Suk Shim, Kenji Sagae, and Louis-Philippe Morency. 2014. Verbal behaviors and persuasiveness in online multimedia content. In *Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP)*, pages 50–58.

Robert B Cialdini and Nathalie Garde. 1987. *Influence*, volume 3. A. Michel.

Domenic V Cicchetti. 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4):284.

Sanorita Dey, Brittany Duff, Karrie Karahalios, and Wai-Tat Fu. 2017. The art and science of persuasion: Not all crowdfunding campaign videos are the same. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 755–769. ACM.

Vincent Etter, Matthias Grossglauser, and Patrick Thiran. 2013. Launch hard or go home!: Predicting the success of kickstarter campaigns. In *Proceedings of the First ACM Conference on Online Social Networks*, COSN '13, pages 177–182, New York, NY, USA. ACM.

Noura Farra, Swapna Somasundaran, and Jill Burstein. 2015. Scoring persuasive essays using opinions and their targets. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 64–74.

Debanjan Ghosh, Aquila Khanam, Yubo Han, and Smaranda Muresan. 2016. Coarse-grained argumentation features for scoring persuasive essays. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 549–554.

Debanjan Ghosh, Smaranda Muresan, Nina Wacholder, Mark Aakhus, and Matthew Mitsui. 2014. Analyzing argumentative discourse units in online interactions. In *Proceedings of the First Workshop on Argumentation Mining*, pages 39–48.

Michael D Greenberg, Bryan Pardo, Karthic Hariharan, and Elizabeth Gerber. 2013. Crowdfunding support tools: predicting success & failure. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pages 1815–1820. ACM.

Ivan Habernal and Iryna Gurevych. 2016. Which argument is more convincing? analyzing and predicting convincingness of web arguments using bidirectional lstm. In *ACL*, pages 1589–1599, Berlin, Germany. Association for Computational Linguistics.

Ivan Habernal and Iryna Gurevych. 2017. Argumentation mining in user-generated web discourse. *Computational Linguistics*, 43(1):125–179.

Christopher Hidey, Elena Musi, Alyssa Hwang, Smaranda Muresan, and Kathy McKeown. 2017. Analyzing the semantic types of claims and premises in an online persuasive forum. In *Proceedings of the 4th Workshop on Argument Mining*, pages 11–21.

Carl I Hovland, Irving L Janis, and Harold H Kelley. 1953. Communication and persuasion; psychological studies of opinion change.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Percy Liang. 2005. Semi-supervised learning for natural language. Master's thesis, Massachusetts Institute of Technology.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Association for Computational Linguistics (ACL) System Demonstrations*, pages 55–60.

Tanushree Mitra and Eric Gilbert. 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*, pages 49–61. ACM.

Elena Musi, Debanjan Ghosh, and Smaranda Muresan. 2018. Changemyview through concessions: Do concessions increase persuasion? *arXiv preprint arXiv:1806.03223*.

Maxime Oquab, Léon Bottou, Ivan Laptev, and Josef Sivic. 2015. Is object localization for free?-weakly-supervised learning with convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 685–694.

Andreas Peldszus and Manfred Stede. 2013. From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31.

Isaac Persing and Vincent Ng. 2017. Why can't you convince me? modeling weaknesses in unpersuasive arguments. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 4082–4088. AAAI Press.

Richard E Petty, John T Cacioppo, and David Schumann. 1983. Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of consumer research*, 10(2):135–146.

Pedro O Pinheiro and Ronan Collobert. 2015. From image-level to pixel-level labeling with convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1713–1721.

Reid Pryzant, Young-Joo Chung, and Dan Jurafsky. 2017. Predicting sales from the language of product descriptions. In *Proceedings of the SIGIR 2017 Workshop on eCommerce (ECOM 17)*.

Reid Pryzant, Kelly Shen, Dan Jurafsky, and Stefan Wagner. 2018a. Deconfounded lexicon induction for interpretable social science. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, volume 1, pages 1615–1625.

Reid Pryzant, Kazoo Sone, and Sugato Basu. 2018b. Interpretable neural architectures for attributing an ad's performance to its writing style. In *EMNLP Workshop on BlackboxNLP*.

Sara Rosenthal and Kathleen Mckeown. 2017. Detecting influencers in multiple online genres. *ACM Trans. Internet Technol.*, 17(2):12:1–12:22.

Christos Sardianos, Ioannis Manousos Katakis, Georgios Petasis, and Vangelis Karkaletsis. 2015. Argument extraction from news. In *Proceedings of the 2nd Workshop on Argumentation Mining*, pages 56–66.

LJ Shrum, Min Liu, Mark Nespoli, and Tina M Lowrey. 2012. Persuasion in the marketplace. *The SAGE Handbook of Persuasion*, page 314.

Christian Stab and Iryna Gurevych. 2014. Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 46–56, Doha, Qatar. Association for Computational Linguistics.

Christian Stab and Iryna Gurevych. 2017. Parsing argumentation structures in persuasive essays. *Computational Linguistics*, 43(3):619–659.

Christian Stab, Tristan Miller, and Iryna Gurevych. 2018. Cross-topic argument mining from heterogeneous sources using attention-based neural networks. *arXiv preprint arXiv:1802.05758*.

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. Winning arguments: Interaction dynamics and persuasion strategies in good-faith online discussions. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 613–624, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Zhongyu Wei, Yang Liu, and Yi Li. 2016. Is this post persuasive? ranking argumentative comments in online forum. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 195–200, Berlin, Germany. Association for Computational Linguistics.

Anbang Xu, Xiao Yang, Huaming Rao, Wai-Tat Fu, Shih-Wen Huang, and Brian P. Bailey. 2014. Show me the money!: An analysis of project updates during crowdfunding campaigns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 591–600, New York, NY, USA. ACM.

Diyi Yang and Robert E Kraut. 2017. Persuading teammates to give: Systematic versus heuristic cues for soliciting loans. *Proceedings of the ACM on Human-Computer Interaction*, 1:114.

Zichao Yang, Zhiting Hu, Ruslan Salakhutdinov, and Taylor Berg-Kirkpatrick. 2017. Improved variational autoencoders for text modeling using dilated convolutions. *arXiv preprint arXiv:1702.08139*.

Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1480–1489, San Diego, California. Association for Computational Linguistics.

Fan Zhang and Diane Litman. 2015. Annotation and classification of argumentative writing revisions. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 133–143.

Fan Zhang, Diane Litman, and Katherine Forbes-Riley. 2016. Inferring discourse relations from pdtb-style discourse labels for argumentative revision classification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 2615–2624.