

Reviews For Paper

Paper ID249

TitleComputing Cross-Cultural Differences in Social Media: A Bilingual Word Embedding Approach

Masked Reviewer ID: Assigned_Reviewer_2

Review:

Question	
[Summary] Please summarize the main claims/contributions of the paper in your own words.	This paper suggests a bilingual word embedding method for social media content, using “social vocabularies”. The authors use the model for two tasks, detecting cross-language differences in the use of named Entities in social media (2) Explaining a slang term from the source language by a set of words in the target language. The authors perform evaluation of the proposed method on those two tasks using data from Twitter and Weibo.
[Relevance] Is this paper relevant to an AI audience?	Likely to be of interest to a large proportion of the community
[Significance] Are the results significant?	Moderately significant
[Novelty] Are the problems or approaches novel?	Somewhat novel or somewhat incremental
[Soundness] Is the paper technically sound?	Technically sound
[Evaluation] Are claims well-supported by theoretical analysis or experimental results?	Somewhat weak
[Clarity] Is the paper well-organized and clearly written?	Satisfactory
[Detailed	<div>Pros:<ul style="list-style-type: none">- Addresses an interesting problem- The method seems to be technically sound- Experiments show some improvement over the baselines (although see the caveat below)</div> <div>Cons<ul style="list-style-type: none">- Limited technical novelty and significance- The evolution results lack confidence intervals, and the advantage over some baselines is not very obvious- Could have used better bilingual embedding baselines.- Organization could improve (e.g., when describing "Building SocVec")</div>

Comments] Please elaborate on your assessments and provide constructive feedback.	<p>This is overall a good paper that addresses an important problem. However, the novelty and depth of technical contribution seems to be limited. The main novelty here seems to be the introduction of the “social vocabularies”. Namely, given a bilingual embedded space Specifically, it’s not very clear whether better bilingual embedding models would not achieve similar accuracy numbers.</p> <p>Also, based on the analysis of Table 1, it seems the proposed approach captures platform-specific (and even time-specific) differences in term usage, rather than substantial cultural differences. For instance, I believe that a topic analysis of all the tweets that are relevant to Maldives would reveal that “holiday, travel, etc” are a relevant topic. The fact that it does not show up in this particular example might be a temporal artifact, or due to limitation of the topic modeling approach. Similarly, there was quite a substantial Twitter coverage of the tragic murders of the USC students from China, and the fact that it does not show up in trending topics can be attributed to temporal variations in the topic distribution.</p>
[QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.	<p>Question/Comments:</p> <ul style="list-style-type: none"> - How are the social vocabularies constructed? - Given the relatively small difference between the proposed method and the Duong-BSL (Table 1), do the authors think those results are robust given non-perfect inter-annotator agreement of 0.672? - Would be interesting to consider more bilingual embedding baselines, e.g., Gouws et. al. BiBOWA: Fast Bilingual Distributed Representations without Word Alignments - The paper needs proofreading, e.g., “Is supposed to be” instead of “is suppose to be”, “named entities” instead of “name entities”
[OVERALL SCORE]	Marginally below threshold
[CONFIDENCE]	Reviewer is knowledgeable in the area

Masked Reviewer ID: Assigned_Reviewer_3

Review:

Question	
[Summary] Please summarize the main claims/contributions of the paper in your own words.	This paper aims at understanding words in different languages based on social information related to the words as opposed to only literal translation.
[Relevance] Is this paper relevant to an AI audience?	Of limited interest to an AI audience
[Significance] Are the results significant?	Moderately significant
[Novelty] Are the problems or approaches novel?	Novel
[Soundness] Is the paper technically sound?	Has minor errors
[Evaluation] Are claims well-	

supported by theoretical analysis or experimental results?	Sufficient
[Clarity] Is the paper well-organized and clearly written?	Good
[Detailed Comments] Please elaborate on your assessments and provide constructive feedback.	I appreciate the long evaluation performed to compare the authors' solution to other approaches. However there are some main technical issues in the paper that are not clear enough. For example what happens if the CSV set remains empty. Furthermore, C_i is defined as a word. What is the meaning of $\max(C_1, \dots, C_N)$. What are the values associated with the words C_i ?
[QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.	How does the pseudo word generator compute the values of the words?
[OVERALL SCORE]	Marginally below threshold
[CONFIDENCE]	Reviewer is knowledgeable but out of the area

Masked Reviewer ID: Assigned_Reviewer_4

Review:

Question	
[Summary] Please summarize the main claims/contributions of the paper in your own words.	This paper introduces a new method for cross lingual word embeddings by encoding them in a third latent space made of social words. The approach uses a combination of translation, word embeddings, and bilingual lexicons to create comparable vectors across languages. These vectors are tested with two evaluations: (1) finding cultural differences and (2) translating slang.
[Relevance] Is this paper relevant to an AI audience?	Relevant to researchers in subareas only
[Significance] Are the results significant?	Moderately significant
[Novelty] Are the problems or approaches novel?	Somewhat novel or somewhat incremental
[Soundness] Is the paper technically sound?	Technically sound
[Evaluation] Are claims well-supported by theoretical analysis or experimental results?	Not convincing

[Clarity] Is the paper well-organized and clearly written?	Poor
[Detailed Comments] Please elaborate on your assessments and provide constructive feedback.	<p>I think the paper is likely technically sound but I found myself greatly confused by the presentation and clarity, which lead me to recommend rejection.</p> <p>First, it is not entirely clear to me how the two tasks are evaluated. For task 1, what does it mean to have cultural differences? The text says a score > 3.0, but in the context of the task, it's difficult to understand how frequent co-occurring words might be rated this way. I did not understand what exactly the model was being tested on. For task 2, is this a translation task? How does the definition play a role in evaluation? Again, critical details are left out. If this is a word-to-word translation task, why would you want to translate a word out of context? Wouldn't this mean that if your system worked well, its vectors would be less useful in the general case?</p> <p>Second, I did not follow why the word meanings are projected into this social vocabulary space. The authors add some explanation on page 2 (top left) with "this projection is supposed to carry social and cultural context such as opinions, sentiments and cognition associated with the terms in respective languages" However, it's difficult to see why the particular set of lexica that were chosen are the right set for bilingual embedding. What if different data were used? How sensitive is the method to differences and overlap between the two vocabularies. The approach needs better motivation.</p>
[QUESTIONS FOR THE AUTHORS] Please provide questions for authors to address during the author feedback period.	<ul style="list-style-type: none"> - What is the input and output to the system for Task 1? - What is the input and output to the system for Task 2? - How sensitive is the method to the choice of the social vocabulary?
[OVERALL SCORE]	Marginally below threshold
[CONFIDENCE]	Reviewer is knowledgeable but out of the area