============================================================================

REVIEWER #1

============================================================================

---------------------------------------------------------------------------

Reviewer's Scores

---------------------------------------------------------------------------

Relevance: 4

Presentation/Clarity of style: 3

Originality/Novelty: 3

Scholarship/References: 4

Technical Soundness: 3

Overall recommendation: 4

---------------------------------------------------------------------------

Comments

---------------------------------------------------------------------------

The paper presents approaches to metaphor detection in a corpus of poems. In

general, the detection of metaphors is not a trivial task. However, I am not

quite sure how the metaphors in poems differ from metaphors in fiction and

other literary works.

**VK : Metaphors in poems are more of a comparative type involving describing things with a metaphorical intent instead of slavish descriptions. (Am I right Prof Chris?)**

The authors rely on rules and statistical approaches for

training classifiers. They chose to focus on noun-centered metaphors. I am also

not sure how they observed that poetry has more noun-centered metaphors in

comparison to verb-centered ones.

**VK : The percentage of metaphors in Type 2 POS tag sequences are far less than Type 1 tag sequences. This is just a preliminary observation.**

An in-house corpus has been built. Five types

of metaphors are proposed, but authors concentrate only on the first one,

presented by the pattern: noun - copular verb - noun.

In the rule-based and the statistical approaches the rules and the features

should be presented in a more compact way.

**VK : We think feature engineering is the most complex aspect of metaphor detection & we have spent the most time on that. Therefore, we have explained it in detail.**

Also, some justification should be

given on the selection of the classifiers.

**VK : Isn’t classification supposed to be hit and trial? We run various classifiers and report results on the top ones.**

I really doubt that the authors will be able to distinguish a poetic type of

metaphor. Maybe this is not necessary.

**VK : It is too early to say something. We are planning to build word embeddings on poetry. Maybe it helps to distinguish poetic metaphors.**

They might focus on the detection

improvement instead and also on the coverage.

**VK : We are working on improving precision by applying deep learning classifiers. And for improving coverage, we are working on generic metaphor detection strategy.**

============================================================================

REVIEWER #2

============================================================================

---------------------------------------------------------------------------

Reviewer's Scores

---------------------------------------------------------------------------

Relevance: 5

Presentation/Clarity of style: 5

Originality/Novelty: 3

Scholarship/References: 5

Technical Soundness: 5

Overall recommendation: 5

---------------------------------------------------------------------------

Comments

---------------------------------------------------------------------------

The authors propose a metaphor detection method for poetry that combines

rule-based and statistical methods. They build an annotated corpus on which to

evaluate their system, and compare to other prior work. The agreement of the

annotators was marginal, but performance of the system is good.

**VK : Agreement on metaphor annotation is always tricky due to so much ambiguity.**

The paper focuses on metaphors in the local context. The authors replace prior

work's emphasis on selection restrictions with word embeddings trained on large

data sets.

The authors categorize metaphor into five types, based on the POS tag

sequences. This strikes me as odd, but I suppose there are justifications for

it in the work presented. I would have liked to see at least a short discussion

of how these "types" of metaphor line up with more literary categorizations, if

they exist.

**VK : All of these 5 types are conceptual metaphors in a literal sense. The POS type categorization is more of a NLP one rather than literary to ease the detection task.**

The authors had independent annotators annotate 720 sentences of containing

type I word sequences, and about half of the sentences contained metaphors of

this type. The kaapa agreement was 0.39, which is on the low side, so I would

have liked to see more effort put into raising this number.

**VK : The two prime annotators were asked to justify their annotations. Even if they still disagreed, the third annotator was there to resolve disagreements.**

Nevertheless, the description of the technique is fairly clear and it seems to

work reasonably well. There are numerous comparisons to other work, and the

authors explore a very large set of different classifiers. The authors release

the data, which is very useful.

============================================================================

REVIEWER #3

============================================================================

---------------------------------------------------------------------------

Reviewer's Scores

---------------------------------------------------------------------------

Relevance: 5

Presentation/Clarity of style: 3

Originality/Novelty: 3

Scholarship/References: 2

Technical Soundness: 4

Overall recommendation: 3

---------------------------------------------------------------------------

Comments

---------------------------------------------------------------------------

This paper proposes and evaluates a method for identifying metaphor in poetry.

Given a noun-verb-noun sequence where the verb is a copula, the method performs

a binary classification: whether the sequence is a metaphor.

The method combines rules based on Concrete-Abstract, Concrete-Class- Overlap

and ConceptNet features, and statistical training with GloVe vectors. A test

set of 487 instances have been manually labelled for evaluation purposes. The

proposed method achieved an F-measure of 0.781 with an SVM, outperforming the

rule-based only approach.

This is an interesting research direction. The methodology seems solid, and

the evaluation is thorough. The paper can nonetheless be improved as follows:

(1) The main gap in this paper is the lack of error analysis. It would be

helpful to explain how/why the proposed method outperforms the rule-based

method, especially the instances where the latter fails to classify correctly.

**VK : Yes. We need to do error analysis.**

(2) It is unclear how the results on the “non-metaphor” and “literal”

classes are relevant to the discussion.

**VK : Though detecting metaphors is our prime task, we cannot ignore the “non-metaphor” class. We need to have an acceptable F-score for that as well to maintain the credibility of our classification.**

(3) Discussion on previous work is currently included in the Introduction

section. It should form its own section, and should be expanded to compare the

proposed method with previous work on measuring semantic relatedness between

words.

**VK : Ok. Semantic relatedness is such an abstract field to cite. We tried to focus on metaphor-only semantic word connections and related works, instead of just going for the crude task.**

(4) Since it is rather subjective to make a judgment on metaphor, more

information on the manual annotation is needed, for example guidelines on how

to decide borderline cases, and the nature of the disagreement between

annotators.

**VK : Ok. We can provide the example guidelines. We have attached the whole annotation set with our submission to give an idea about the sentences and their annotations.**

(5) It would be informative to report statistics on the distribution of the 5

types of metaphor.

**VK : Good suggestion. We have not started with the other types. Instead we started working on the generic approach to encompass all the types.**