

# Authors' responses to reviewer's comments

The authors thank the reviewers for their valuable feedbacks, which have further improved the presentation of the extended version. In this document, we provide responses to the comments raised by reviewers.

Best regards,

M. G. Carneiro, T. H. Cupertino, L. Zhao and J. L. G. Rosa

## 1 Reviewer 1

**C-1.1:** *The paper discusses the application of semi-supervised learning for classifying arguments in the context of Semantic Role Labeling (SRL). The paper is properly motivated as in the domain there is plenty of unlabeled data and a reduced number of labeled predicates. The experimental evaluation is complete and shows the approach soundness.*

**A-1.1:** Thanks for your positive comments.

**C-1.2:** *However, the discussion of the experimental findings is short and oriented to discuss the algorithms parameters. I believe the results should be contrasted with the (probably poor) results of supervised classification with a reduced number of examples as well as show how semi-supervised learning can evolve as more examples are labeled.*

**A-1.2:** Authors thank the reviewer for this comment. We have included both analysis at the extended version.

**C-1.3:** *Also, it would be good to explain why kNN was selected instead of classical semi-supervised approaches as EM, which is grounded on the probability distributions of the unlabeled data.*

**A-1.3:** Thanks for this comment. It is a good opportunity to say we tested two techniques using EM for Gaussian mixture models, however the results were very poor. We believe this is because the classes does not correspond to groups. In addition, our simulations employs very few labeled data (e.g., one per class) which also makes difficult to EM-GMM generating an appropriate a priori probability from the labeled data.

**C-1.4:** *A minor comment is that the example within the introduction seems like out of place, it would be better to explain that on section 2 perhaps.*

**A-1.4:** Authors thank the reviewer. The comment has been addressed at the extended version.

## 2 Reviewer 2

**C-2.1:** *The main purpose of the work described in this paper was to evaluate the application of graph-based SSL methods for semantic role labeling on a Brazilian Portuguese corpus named PropBank-br. The subject addressed in the paper is not new, but it is relevant and it is within the scope of the Symposium. The paper is well written and well organized. The experimental methodology adopted seems appropriate and correct.*

**A-2.1:** Thanks for your positive comments.

**C-2.2:** *However, the paper lacks of a proper discussion about the results obtained in the experiments. The authors merely describe what is already possible to see in the graphs and tables (for example, which algorithm had the best performance). If the authors had presented an explanation about the reasons for the results achieved it certainly would have enriched the paper.*

**A-2.2:** We thank the reviewer for this comment. At the extended version, authors included additional analysis which better explain some results achieved, e.g.:

*Although self-training is a simple semi-supervised heuristic, the experimental results presented in this article have shown such method is very promising for Brazilian Portuguese semantic role labeling. The usage of logistic regression as a base classifier in order to learn the unlabeled data iteratively from a very few labeled data allows good predictive performance. In addition, common problems related to self-training, such as the inclusion of errors by labeling unlabeled data wrongly, does not affect the learning process as we can see by examining the execution step by step. A probable reason is the nature of the data which is too sparse and the simplicity of the logistic regression which avoids overfitting.*

## 3 Reviewer 3

**C-3.1:** *The paper presents a strategy for semi-supervised learning of semantic roles. The strategy is based on the analysis of the adjacency matrix built from the text. Results on a benchmark dataset are reasonable and improve to some extent considering existing techniques.*

**A-3.1:** Thanks for your positive comments.

**C-3.2:** *My first criticism to the paper is that it is not clear its innovation in terms of the technique, since the strategies employed have been previously proposed. I suggest the authors to make clear the novelty of their work.*

**A-3.2:** Authors thank the reviewer for this comment. It is a good opportunity to clarify that this article represents the first attempts to employ semi-supervised learning for Brazilian Portuguese semantic role labeling. The scarcity of annotated data in the PropBank-br corpus indicates an appropriate scenario to investigate machine learning techniques able to learn not only from labeled data, but specially from the unlabeled ones. However, to the best of our knowledge, there is no such kind of publication for Brazilian Portuguese language in

literature. The following paragraph has been added at the extended version in order to address the reviewer comment.

*“By contrast, to the best of our knowledge, semi-supervised learning is also a barely explored topic for Brazilian Portuguese SRL, whose the main work, which investigated the propagation of semantic roles under a graph-based semi-supervised framework, is presented in [Carneiro et al. 2016]. Here, we extend that work by investigating not only graph-based SSL methods, but also other categories of SSL, such as low-density separation and self-training.”*

**C-3.3:** *Second, the although apparently sound, the rationale behind the various steps proposed should be presented. As it is, it assumes that the audience understands all the concepts the paper inherits from previous work and it is not clear how it is going to perform the targetted task.*

**A-3.3:** Authors thank the reviewer for the comment. The rationale behind the main steps of the work, including the preprocessing of PropBank-br, are described in more details at the extended version.

**C-3.4:** *Third, in the experiments, I suggest the authors to perform some statistical significance tests on the results presented in Table II, since the standard deviations are pretty high and being the best average is not a guarantee that the result really outperforms the other techniques and parameter configurations.*

**A-3.4:** We thank the reviewer for the suggestion. Statistical tests have been included at the extended version,

**C-3.5:** *Overall, the paper is well written, although it has some typos and grammar errors that must be fixed. I suggest the authors to perform a careful review prior to submit a camera-ready version. For instance, in the second to last paragraph of the intro, the paper says “This give us...”, where the verb is lacking an “s” at the end and the expression is not reading well. Please rephrase.*

**A-3.5:** Authors thank the reviewer by providing such corrections.