Review Form

Submission #343: XRCE at SemEval-2016 Task 5: Feedbacked Ensemble Modeling on Syntactico-Semantic Knowledge for Aspect Based Sentiment Analysis **Submission Type:** System Description **Task:** Task 5: Aspect-Based Sentiment Analysis	**Reviewer:** Zhiqiang Toh **Secondary Reviewer:**

Summary Ranking

Please evaluate the submission according to the criteria below.

Evaluation Category	Enter Your Score
Appropriateness Is this paper appropriate for SemEval? It should be a system description paper for a participating team.	○ Appropriate (most submissions) ○ Not appropriate
Clarity (1-5) For a reasonably well-prepared reader, is it clear what was done and why? Is the paper well-written and well-structured? 5 = Very clear. 4 = Understandable by most readers. 3 = Mostly understandable with some effort. 2 = Important questions were hard to resolve even with effort. 1 = Much of the paper is confusing.	○ 1 ○ 2 ○ 3 ◉ 4 ○ 5
Soundness (1-5) Is the technical approach sound and well-chosen? Can one trust the claims of the paper – are they supported by proofs or proper experiments where the results of the experiments are correctly interpreted? 5 = The approach is sound, and the claims are convincingly supported. 4 = Generally solid, but there are some aspects of the approach or evaluation I am not sure about. 3 = Fairly reasonable, but the main claims cannot be accepted based on the material provided. 2 = Troublesome. Some interesting ideas, but the work needs better justification or evaluation. 1 = Fatally flawed.	○ 1 ○ 2 ○ 3 ◉ 4 ○ 5

Replicability (1-5) Will members of the research community be able to reproduce or verify the results described in this paper? A lower score might be assigned if an insufficient amount of detail has been provided, if there is a highly subjective component to the setting of certain parameters, or if proprietary data have been used in the experiments. A low score here does not necessarily imply a low overall recommendation. Members of the SemEval community, and of the NLP community in general… 5 = could easily reproduce the results and verify the correctness of the results described here. 4 = could mostly reproduce the results described here, although there may be some variation because of sample variance or minor variations in their interpretation of the protocol or method. 3 = could possibly reproduce the results described here with some difficulty. The settings of parameters are underspecified or very subjectively determined; the training data required are not widely available. 2 = could not reproduce the results described here no matter how hard they tried. The author simply has not provided a sufficient amount of detail nor access to resources for us to do anything more than accept their conclusions without question. 1 = not applicable (please use this very sparingly, such as for short submissions that are opinion pieces).	○ 1 ○ 2 ◉ 3 ○ 4 ○ 5
Overall Recommendation (Old) In deciding on your ultimate recommendation, please think over all your scores above, but remember that no paper is perfect, and that SemEval has traditionally accepted all system description papers that have been properly written; the reject rate has been extremely low. Remember that the author has about three weeks to address reviewer comments before the camera-ready deadline.	◉ Accept ○ Reject
"Best Of" Recommendation This system is highly novel, insightful, or impactful and deserves to be in the "Best Of" session.	○ No ◉ Maybe ○ Yes

Detailed Comments

Please supply detailed comments to back up your rankings. These comments will be forwarded to the authors of the paper. The comments will help the committee decide the outcome of the paper, and will help justify this decision for the authors. Moreover, if the paper is accepted, the comments should guide the authors in making revisions for a final manuscript. Hence, the more detailed you make your comments, the more useful your review will be - both for the committee and for the authors.

```
This paper describes the work on Subtask 1 of ABSA for the English and
French
datasets. An in-house parser is used to extract syntactic and semantic
features
for domain term detection, aspect category classification and polarity
classification. Domain term detection uses CRF as the training algorithm,
while
feedbacked ensemble models are used for the classification tasks.
Evaluation
results suggest that the proposed framework is feasible for the task,
especially polarity classification, which is ranked 1st for both English
and
French.

The paper is generally well-written. However, more details about how the
feedback is used to adjust feature selection can be included to benefit the
readers. E.g. whether the process requires manual supervision.

Some grammatical errors:

whit -> with

adress -> address
```

Confidential Comments for Committee

You may wish to withhold some comments from the authors, and include them solely for the committee's internal use. For example, you may want to express a very strong (negative) opinion on the paper, which might offend the authors in some way. Or, perhaps you wish to write something which would expose your identity to the authors. If you wish to share comments of this nature with the committee, this is the place to put them.

Submit
