

Summary

- We assess the performance of current unsupervised aspect-based opinion mining techniques to a new domain of text – Feedback Surveys
- We make improvements to existing algorithms by the following approaches:
 - (1) Optimize aspect extraction for precision rather than recall, but still maintain a high level recall for *important* aspects
 - (2) Handling language features that are common in the feedback survey domain (modals)

Motivation

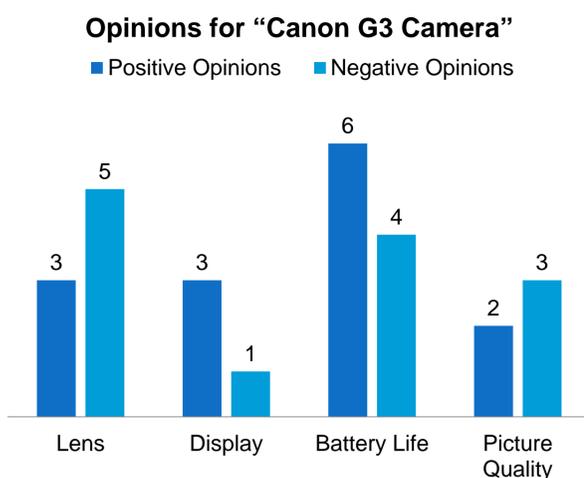
Opinion Mining

- The influence of social media and the rapid expansion of e-commerce has caused great interest in opinion mining. Opinion mining allows for the analysis and summarization of high volumes of user opinions by extracting *structured* opinion information from *unstructured* text.
- There have been many unsupervised methods proposed for opinion mining. Using unsupervised methods is advantageous because they do not required training data.
- However, very little research has been performed in assessing and their performance in other domains e.g Feedback Surveys.
- There are also many ways that current methods can be improved to perform better for the general case, as well as for the Feedback Survey domain

Background

Aspect-based Opinion Mining

- Aspect-based opinion mining is the extraction of opinions expressed by an opinion *source* towards a specific *aspect* of a known *topic*.
- The goal is to extract (source, aspect, orientation) triplets for a collection of user comments in a given topic, and summarize it in structured a way similar to below:



Goals and Contributions

- Evaluation in the current literature is very subjective. We address this issue by defining a **rigorous and objective evaluation framework** (see thesis)
- We apply and evaluate the performance of existing unsupervised methods to a **new domain – Feedback Surveys**
- We **improve current techniques** with the goal of increasing accuracy for both the general case, and for the feedback survey domain

Existing Methods

Aspect Extraction

- The existing method (Qiu et al. 2011) chosen utilizes dependency relation between aspects and opinion words to extract new aspects. There are two main rules:
 - Extract Aspects using known opinion words
 - Extract new aspects using known extracted aspects via a bootstrapping process

Orientation Prediction

- The existing method (Ding et al. 2008) utilizes an *Opinion Lexicon* and a set of rules to handle the effects of *negation* and *context-dependency* on the orientation of modifying opinion words.
- It then summarizes the opinion words which are modifying a target aspect using the scoring function:

$$score(a) = \sum_{w_i} \frac{orientation(w_i)}{dist(w_i, a)}$$

Proposed Improvements

General Improvements

- For aspect extraction we propose to extract candidate aspects that strictly co-occur with *aspect indicators*. This is when:
 - The candidate aspect is modified by an opinion word, or;
 - The candidate aspect is used in conjunction with a known extracted aspect

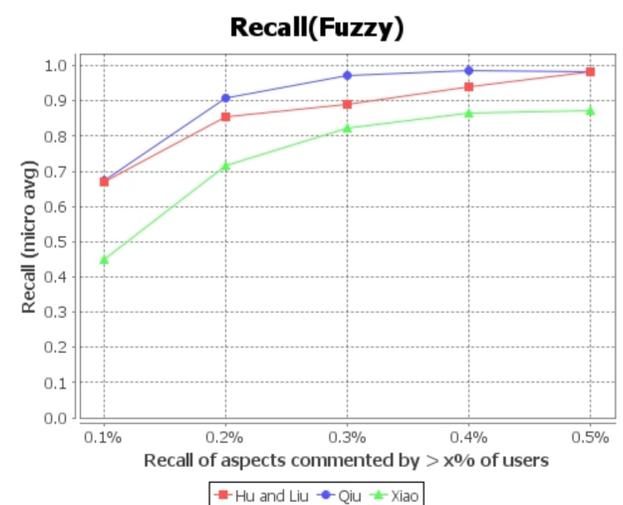
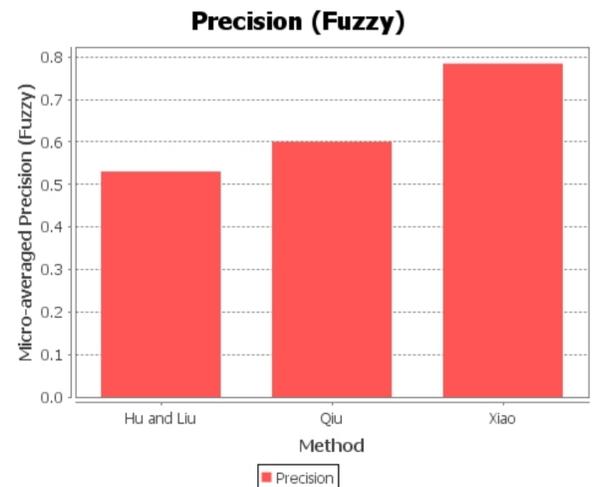
Feedback Surveys Improvements

- Feedback surveys contain many *recommendations* by users to improve a particular aspect. These are indicated by *modal verbs* (should, could etc)
- We utilize this common language feature in feedback surveys to improve accuracy by:
 - (1) Define new dependency relation patterns for modals to extract aspects targeted by recommendation opinions
 - (2) Define new rules to handle modals for orientation prediction

Results

Aspect Extraction

- ‘Xiao’ is the proposed technique



Orientation Prediction

- ‘Xiao’ is the proposed technique

Precision	Ding	Xiao	Increase
Average	0.844	0.873	+3.47%

Recall	Ding	Xiao	Increase
Average	0.868	0.885	+1.90%

Conclusion

Feedback Surveys

- Existing opinion mining methods produce similar performance for the feedback survey domain. However, there are language features (modals) in feedback surveys that can be exploited to improve opinion mining.

Performance of proposed Improvements

- Using *aspect indicators*, precision was increased significantly. Recall is reduced, but recall for important (frequently opinionated) aspects is still high.
- By handling *modal verbs*, precision and recall of orientation prediction and aspect extraction were increased.

References

- Ding, X., Liu, B., Yu, P. 2008. A Holistic Lexicon-Based Approach to Opinion.
- Qiu, G., Liu, B., Bu, J., Chen, C. 2011 Opinion Word Expansion and Target Extraction through Double Propagation.