

Aspect Terms

The lasagna was great, but we had to wait 20 minutes just to be seated.

Aspect term: lasagna (positive sentiment)

Subtask 1: Aspect Term Extraction

Task: to detect aspect terms in a sentence

Approach:

- Semi-Markov discriminative tagger, trained with MIRA
- tags phrases, not tokens, can use phrase-level features

Features:

- emission features: token identity (cased, lowercased) in a 2-word window, prefixes and suffixes up to 3 chars, phrase identity (cased, lowercased)
- transition features: tag ngrams

Results:

Domain	P	R	F1
Restaurants	84.41	76.37	80.19 (3 rd among 24 teams)
Laptops	78.77	60.70	68.57 (3 rd among 24 teams)

Subtask 2: Aspect Term Polarity

Task: to detect sentiment towards a given aspect term

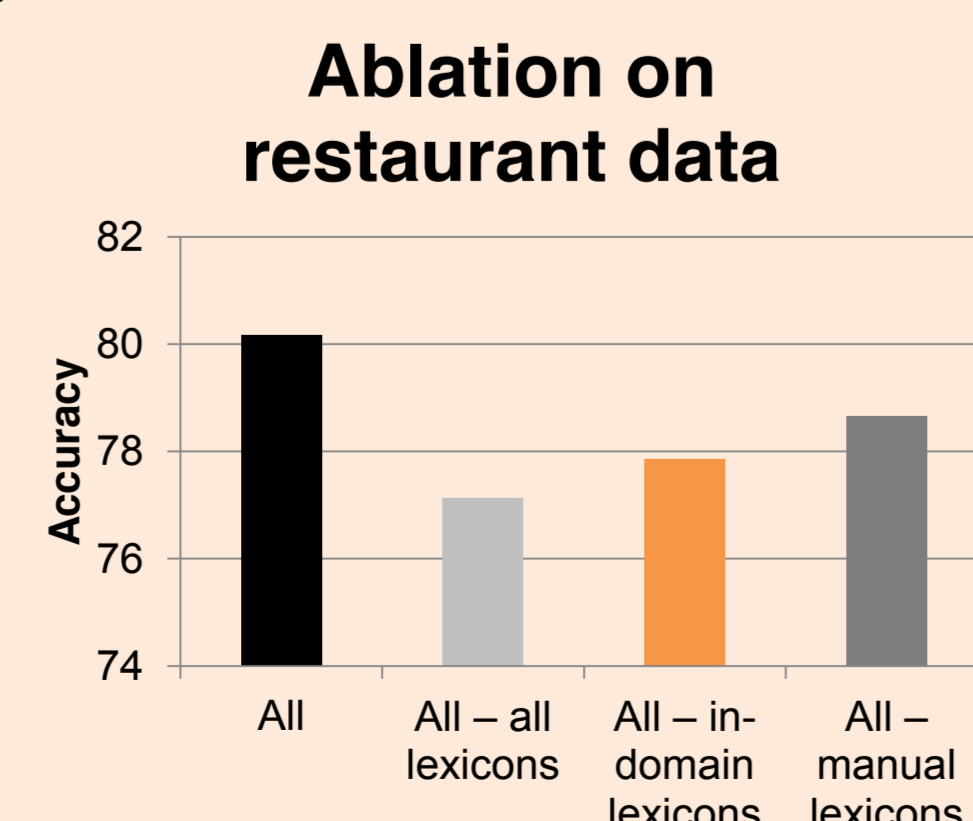
Approach: SVM with linear kernel

Features:

- surface features: ngrams, context-target bigrams
- sentiment lexicon features: counts, sum, max
- syntactic features: ngrams and context-target bigrams on parse trees, parse label features

Results:

Domain	Accuracy
Restaurants	80.16 (2 nd among 29 teams)
Laptops	70.49 (1 st among 29 teams)



We present a sentiment analysis system to detect aspect terms, aspect categories and sentiment expressed towards aspect terms and categories in customer reviews.

Rank obtained by NRC-Canada in SemEval-2014 Task 4

Term: 3 Category: 1 Term Sentiment: 1,2 Cat. Sentiment: 1

- builds on the NRC-Canada sentiment analysis system which determines the overall sentiment of a message (top results in SemEval-2013 Task 2 and SemEval-2014 Task 9 on Sentiment Analysis of Tweets)

Resources

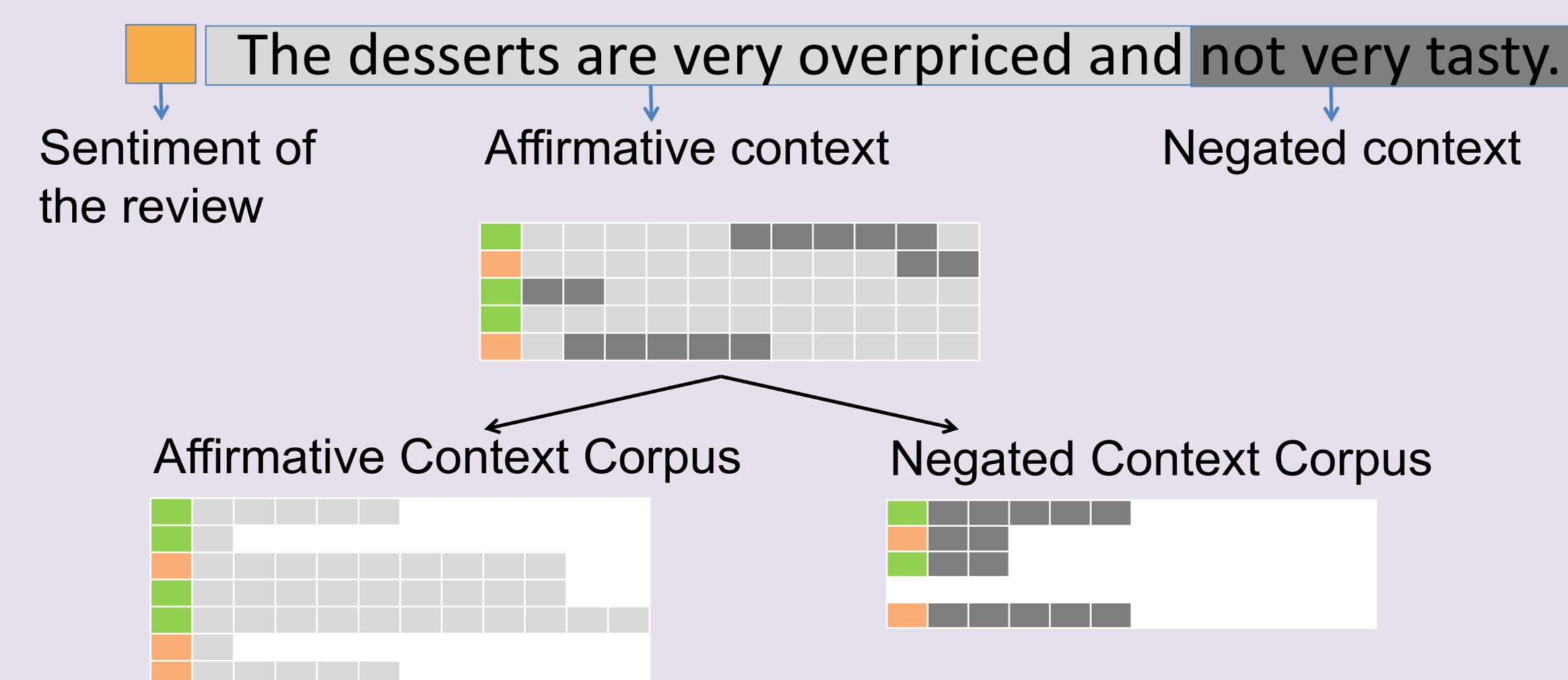
In-domain corpora

- 180,000 Yelp restaurant reviews (Phoenix Academic dataset)
- 125,000 Amazon laptop reviews (McAuley & Leskovec, 2013)

Sentiment lexicons - terms and degree of their association with positive or negative sentiment

Generation process:

1. star ratings in reviews are used as weak labels
2. $score(w) = PMI(w, positive) - PMI(w, negative)$
 - if $score(w) > 0$, then word w is positive
 - if $score(w) < 0$, then word w is negative
3. affirmative and negated contexts are treated separately



Word-Aspect Association lexicon - terms and degree of their association with the aspect categories

Conclusions

- top results on subtasks 2, 3, and 4
- statistical approaches with surface-form and lexicon feat.
- most useful features: derived from automatically generated in-domain lexical resources
- resources to download: www.purl.com/net/sentimentoftweets

Aspect Categories

The lasagna was great, but we had to wait 20 minutes just to be seated.

Aspect categories: food (positive sentiment), service (negative sentiment)

Subtask 3: Aspect Category Detection

Task: to detect aspect categories discussed in a sentence

Approach:

- SVM with linear kernel
- five binary classifiers (one-vs-all)
- assign $c_{max} = \text{argmax}_c P(c|d)$ if $P(c_{max}|d) \geq 0.4$

Features:

- word and character ngrams
- stemmed ngrams (Porter stemmer)
- word cluster ngrams (Brown clustering algorithm)
- Yelp Restaurant Word-Aspect Association lexicon

Results:

Domain	P	R	F1
Restaurants	91.04	86.24	88.58 (1 st among 18 teams)

Subtask 4: Aspect Category Polarity

Task: to detect sentiment towards a given aspect category

Approach:

- one 4-class SVM classifier with 2 copies of each feature: generic and category-specific
- add features for terms associated with aspect category

Features:

- word and character ngrams, POS tags
- word cluster ngrams
- sentiment lexicon features

Results:

Domain	Accuracy
Restaurants	82.93 (1 st among 20 teams)

