**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:**

<table>
<thead>
<tr>
<th>Domain</th>
<th>P</th>
<th>R</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants</td>
<td>84.41</td>
<td>76.37</td>
<td>80.19 (3rd among 24 teams)</td>
</tr>
<tr>
<td>Laptops</td>
<td>78.77</td>
<td>60.70</td>
<td>68.57 (3rd among 24 teams)</td>
</tr>
</tbody>
</table>

---

**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:**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants</td>
<td>80.16 (2nd among 29 teams)</td>
</tr>
<tr>
<td>Laptops</td>
<td>70.49 (1st among 29 teams)</td>
</tr>
</tbody>
</table>

---

**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 $c_{max} = \arg\max_c P(c|d)$ 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:**

<table>
<thead>
<tr>
<th>Domain</th>
<th>P</th>
<th>R</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants</td>
<td>91.04</td>
<td>86.24</td>
<td>88.58 (1st among 18 teams)</td>
</tr>
</tbody>
</table>

---

**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:**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restaurants</td>
<td>82.93 (1st among 20 teams)</td>
</tr>
</tbody>
</table>

---

**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

<table>
<thead>
<tr>
<th>Term: 3</th>
<th>Category: 1</th>
<th>Term Sentiment: 1.2</th>
<th>Cat. Sentiment: 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>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)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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/sentimetoftweets](http://www.purl.com/net/sentimetoftweets)