Continuous Models of Affect from Text using N-Grams

Nikolaos Malandrakis, Alexandros Potamianos, Shrikanth Narayanan
malandra@usc.edu, potam@telecom.tuc.gr, shri@sipi.usc.edu

Introduction
- Creation of continuous affective ratings
  - Words/Terms
  - Sentences
- Compositional assumption
  - Hierarchical decomposition
- Multi-word terms not handled
  - “in short”
  - “look up”
  - “kick the bucket”
- Our approach:
  - Language modeling inspired
  - Bigram terms
  - Back-off to unigrams

Word/Term model
- Ratings through semantic similarities to known words
  \[ \hat{v}(w_j) = a_0 + \sum_{i=1}^{N} a_i v(w_i) d_{ij}, \]
- \( d_{ij} \) cosine similarity of binary weighted context vectors
  - 116m sentence web corpus
- Affective Norms for English Words (ANEW)
  - 1034 annotated words
  - Extrema \( \rightarrow \) semantic space
  - Used to train \( a_i \)

Evaluation
- SemEval’2007 corpus
  - 1000 news headlines
  - Continuous valence
  - 53% negative
  - Train set of 250 headlines
- Binary polarity classification
  - 1grams only > 2grams only
  - Significant improvement
  - Semantic criterion performs best
  - Optimal performance at 75% rejection

Sentence Model
- Tokenization
  - POS Tagging
  - 2-word overlapping windows
    \[ \text{2-word window} \]
- Lexicon Lookup
  \[ v_b(w_2 w_3) = \begin{cases} b_1 & \text{if } 1 \leq t \leq c(1,2) \\ b_2 v(w_1 w_2) & \text{if } c(1,2) > t \end{cases} \]
- Term Selection
  - Use bigram term or back-off to unigrams
  - Criteria of non-compositionality
    - Affective:
      \[ c_a(i,j) = |v(w_i w_j) - 0.5[v(w_i) + v(w_j)]| \]
    - Semantic:
      \[ c_s(i,j) = p(w_i) p(w_j) \log \frac{p(w_i, w_j)}{p(w_i) p(w_j)} \]
- Term Fusion
  \[ v_w(s) = b_0 + \frac{1}{N} \left[ \frac{b_1}{2} (v(w_1) + v(w_N)) + \sum_{i=1}^{N-1} v(w_i w_{i+1}) \right] \]

Sentence Rating

Statistics
- Annotated Lexicon

<table>
<thead>
<tr>
<th></th>
<th>V</th>
<th>A</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>love</td>
<td>0.98</td>
<td>0.36</td>
<td>0.53</td>
</tr>
<tr>
<td>hate</td>
<td>0.72</td>
<td>0.49</td>
<td>0.01</td>
</tr>
<tr>
<td>bored</td>
<td>-0.51</td>
<td>-0.54</td>
<td>-0.22</td>
</tr>
</tbody>
</table>

Conclusions
- Significant improvement over unigrams
- Adaptable compositional frameworks
- Future work:
  - Improved term model
  - Higher order terms
  - Alternate selection criteria

Acknowledgments
- Most of this work performed while N. Malandrakis was with the Dept. of ECE, TU Crete
- Partially supported by the IST Programme of the EU under contract number 296170 (PortDial project)
- Partially funded by the Viterbi Fellowship and NSF