USC University of Southern California

USC Viterbi School of Engineering

Ming Hsieh Department of Electrical Engineering

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

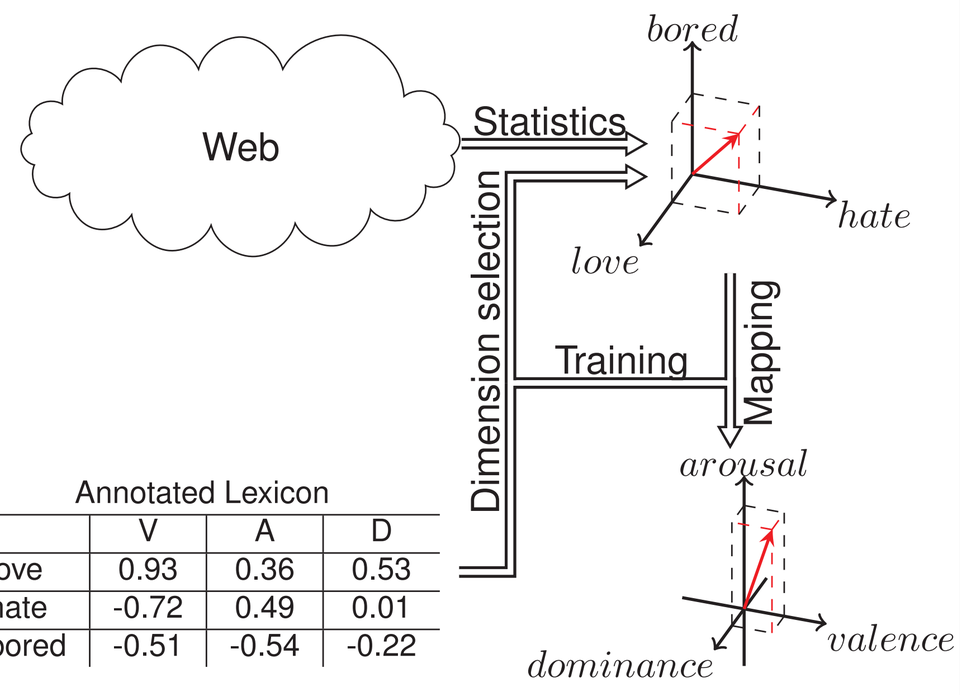
- Creation of continuous affective ratings
  - Words/Terms
  - Sentences
- Compositionality assumption
  - Hierarchical decomposition
- Multi-word terms not handled
  - *"in short"*
  - *"look up"*
  - *"kick the bucket"*
- **Our approach**:
  - Language modeling inspired
  - Bigram terms
  - Back-off 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



Annotated Lexicon

	V	A	D
love	0.93	0.36	0.53
hate	-0.72	0.49	0.01
bored	-0.51	-0.54	-0.22

Sentence Model

Sentence

Tokenization

- POS Tagging
- 2-word overlapping windows

$\underbrace{w_1 \; w_2}_{\text{2-word window}} \; w_3 \; w_4 \; w_5 \; w_6 \; \ldots$

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: $c_s(i,j) = p(w_i w_j) \; \log \frac{p(w_i w_j)}{p(w_i)p(w_j)}$

Lexicon Lookup

$$v_b(w_1 w_2) = \left\{ \begin{array}{ll} b_1 \; \frac{1}{2} \sum_{i=1}^{2} v(w_i), & c(1,2) \leq t \\ b_2 \; v(w_1 w_2), & c(1,2) > t \end{array} \right\}$$

Term Fusion

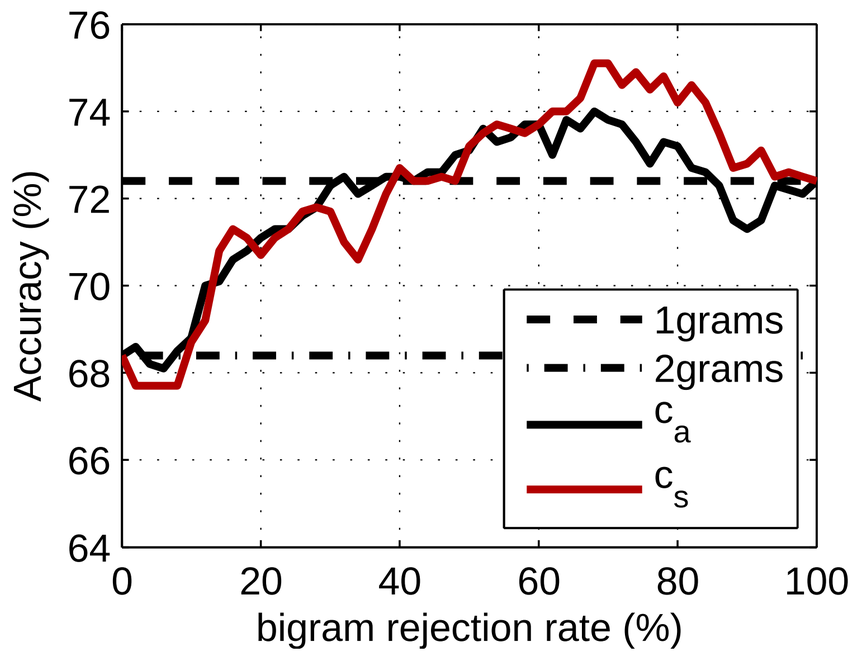
- Linear model trained with LSE

$$v_{bo}(s) = b_0 + \frac{1}{N}\left[\frac{b_1}{2}(v(w_1)+v(w_N)) + \sum_{i=1}^{N-1} v_b(w_i w_{i+1})\right]$$

Sentence Rating

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



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