Supervised Learning of Word Sentiment

Natalie Parde
UIC CS 421

Normalized Word Likelihood

- **Document-level sentiment classifier** → any statistical or neural methods we've learned about so far!
- **Word-level sentiment classifier** → also consider simple probabilistic measures
- **Normalized word likelihood**
  - $P(w|c) = \frac{\text{count}(w,c)}{\sum_{w \in C} \text{count}(w,c)}$

Potts Diagrams

- Mechanism for visualizing word sentiment
  - Sentiment class vs. normalized word likelihood
- Characteristic patterns:
  - **“J” shape:** Strongly positive word
  - **Reverse “J” shape:** Strongly negative word
  - **“Hump” shape:** Weakly positive or negative word
- Patterns may also correspond to different types of word classes
  - Emphatic and attenuating adverbs

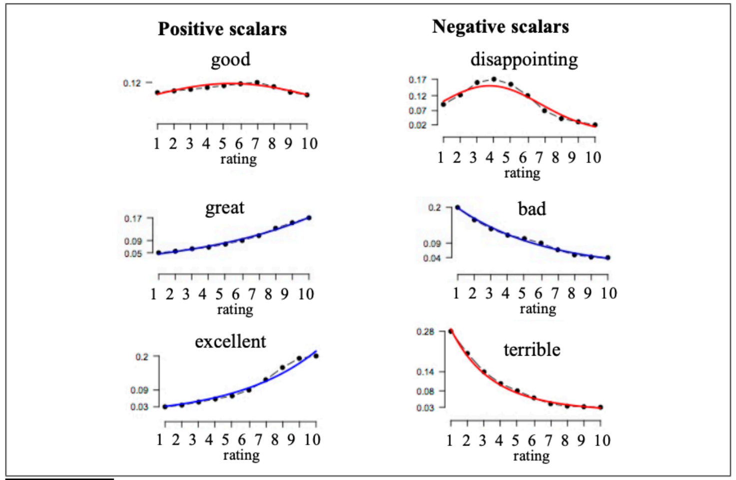

Figure 21.10 Potts diagrams (Potts, 2011) for positive and negative scalar adjectives, showing the J-shape and reverse J-shape for strongly positive and negative adjectives, and the hump-shape for more weakly polarized adjectives.

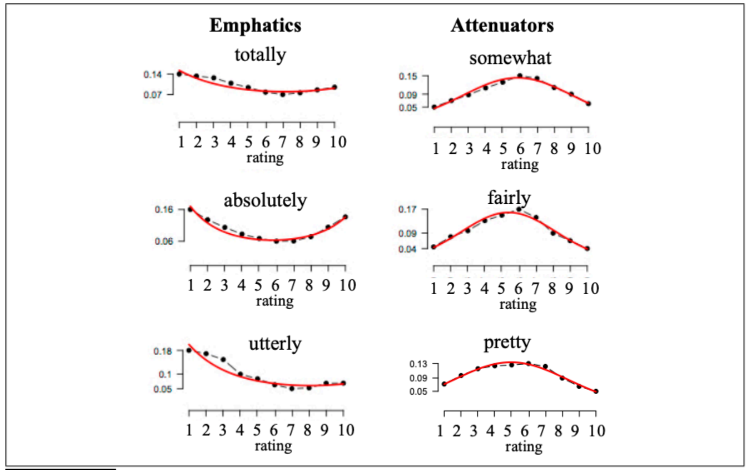

Figure 21.11 Potts diagrams (Potts, 2011) for emphatic and attenuating adverbs.

Log Odds Ratio with Informative Dirichlet Prior

- Is computing raw difference in frequency sufficient for identifying whether a word is very positive or very negative?
  - Not necessarily!
  - **Highly frequent words may have large raw differences even with small relative differences**, and highly infrequent words may have small raw differences even with large relative differences
- More sophisticated approach: **Log odds ratio with an informative Dirichlet prior**

Log Odds Ratio

- **Probability of word *w* existing in corpus *i*:**
  - $P^i(w) = \frac{f_w^i}{n^i}$
- **Log odds ratio:**
  - $\text{lor}(w) = \log \frac{P^i(w)}{1-P^i(w)} - \log \frac{P^j(w)}{1-P^j(w)}$
  - $\text{lor}(w) = \log \frac{f_w^i}{n^i-f_w^i} - \log \frac{f_w^j}{n^j-f_w^j}$

Prior-Modified Log Odds Ratio

- Modifying the previous equation with an informative Dirichlet prior:
  - $\delta_w^{(i-j)} = \log \frac{f_w^i + \alpha_w}{n^i + \alpha_0 - (f_w^i + \alpha_w)} - \log \frac{f_w^j + \alpha_w}{n^j + \alpha_0 - (f_w^j + \alpha_w)}$

Log Odds Ratio with Informative Dirichlet Prior

- Estimate of variance for the modified log odds ratio:
  - $\sigma^2\left(\hat{\delta}_w^{(i-j)}\right) \approx \frac{1}{f_w^i+\alpha_w} + \frac{1}{f_w^j+\alpha_w}$
- Final equation:
  - $\frac{\hat{\delta}_w^{(i-j)}}{\sqrt{\sigma^2\left(\hat{\delta}_w^{(i-j)}\right)}}$