Approaches to Sentiment Analysis

MSE 2400  EaLiCaRA
Spring 2015      Dr. Tom Way

Based in part on notes from Aditya Joshi

What is Sentiment Analysis?

• Identify the orientation of opinion in a piece of text
  - The movie was fabulous!
  - The movie stars Mr. X
  - The movie was horrible!

• Can be generalized to a wider set of emotions

Motivation

• Knowing sentiment is a very natural ability of a human being.
  Can a machine be trained to do it?

• SA aims at getting sentiment-related knowledge especially from the huge amount of information on the internet

• Can be generally used to understand opinion in a set of documents
Challenges

- Contrasts with standard text-based categorization
- Domain dependent
- Sarcasm
- Thwarted expressions

Example: The perfume is so amazing that I suggest you wear it with your windows shut.

The sentences/words that contradict the overall sentiment of the set are in majority.

Example: The actors are good, the music is brilliant and appealing. Yet, the movie fails to strike a chord.

Quantifying sentiment

Each term has a Positive, Negative and Objective score. The scores sum to one.
**Approach 1: Using adjectives**

- Many adjectives have high sentiment value
  - A ‘beautiful’ bag 😊
  - A ‘wooden’ bench 😞
  - An ‘embarrassing’ performance 😞
- Focusing on adjectives might be beneficial

**Approach 2: Using Adverb-Adjective Combinations (AACs)**

- Calculate sentiment value based on the effect of adverbs on adjectives

  **Linguistic ideas:**
  - **Adverbs of affirmation:** certainly
  - **Adverbs of doubt:** possibly
  - **Strong intensifying adverbs:** extremely
  - **Weak intensifying adverbs:** scarcely
  - **Negation and Minimizers:** never

**Approach 3: Subject-based SA**

- Examples:

  The horse bolted. 😞

  The movie lacks a good story. 😞
Lexical Analysis

subj. bolt

Argument that receives the sentiment (subj./obj.)

subj. lack obj.

Argument that sends the sentiment (subj./obj.)

Example

The movie lacks a good story.

The movie lacks \S+.

Lexicon:
- G JJ good obj.
- B VB lack obj ~subj.

Steps:
1) Consider a context window of up to five words
2) Shallow parse the sentence
3) Step-by-step calculate the sentiment value based on lexicon and by adding \"S+\" characters at each step

Applications

- Review-related analysis
- Developing ‘hate mail filters’ analogous to ‘spam mail filters’
- Question-answering (Opinion-oriented questions may involve different treatment)