E-Ship: Modeling public and entrepreneur mood from longitudinal online data

Johan Bollen
School of Informatics, Engineering, and Computing
Approach

- Measure population-wide sentiment over time and validating it against major events, regional indicators, and entrepreneurial activity and resilience.

- E-Ship relevant work:
  - Measuring collective sentiment from online data (county-level, US)
  - Cross-validation against major events and disasters (possibility of measuring resilience, time to recovery)
  - Cross-validation against large-surveys of Subjective Well-Being at state-level in the US
Common Social Media Indicators

ATTENTION:

- Clicks, Likes, Shares, Friends, Brand mentions, Profile visitors, Active followers, Ratings, Fans, Check ins, Time

CONTENT-DERIVED INDICATORS OF PSYCHOLOGY and CULTURE:

- Sentiment
- Personality
- Social indicators (unrest)
- Culturomics

jbollen@Indiana.edu
Sentiment Analysis:

Mapping text content to emotional classification or ratings

- **Polarity rating:**
  - Numerical expression of valence based on content features
  - $F(\text{text}) \rightarrow [-1,+1]$

- **Classification:**
  - Content features: words, terms, etc
  - Discrete positive, negative, or neutral classes

- **Tools:**
  - Python NLTK, GATE, Stanford Sentiment Analysis Module, TextBlob, VADER
VADER - Valence Aware Dictionary and sEntiment Reasoner

- What is valence?
- “gold-standard sentiment lexicon” specifically built for social media
- Gathered words from sentiment word-banks
- Used crowd-sourcing to find sentiment values from [-1,1]
- 5 grammatical and syntactical rules

Social media specific capabilities

1. **Punctuation** (!) modifies intensity
   a. That was great vs. That was great!!!!
2. **Capitalization** (all caps) modifies when present with non capitalized words
   a. That was great vs That was GREAT
3. **Degree Modifiers** (adjectives/ adverbs)
   That was great vs. That was really great
4. **Contrastive Conjunction** (but) forces the latter half to be more dominant
   a. That was fun but I didn’t like it
5. **Trigram analysis** to find negation
   a. That was not that great

VADER - Example

- Valence - intrinsic goodness or badness of a word or phrase
- Compound score - normalized sum of the valence of each word after the rules have been applied
- Positive, neutral, negative - proportion of phrase that falls into each category

"Johan is smart, handsome, and funny."
- Neg: 0.0
- Neu: 0.254
- Pos: 0.746
- Compound: 0.8316

"Today SUX!"
- Neg: 0.779
- Neu: 0.221
- Pos: 0.0
- Compound: -0.5461
Diagram of Data and Processing

Twitter data ("Gardenhose"): random 10%, +/-6 years → Longitude/latitude

140 chars → VADER sentiment analysis: [-1, +1] × time

Geo-aware socio-cultural resource

 jbollen@Indiana.edu
Our data

- **3,221 US counties** - US Census Bureau 2010
- Mean 3434.8 Tweets/county

Distribution follows population density 80/20 law!
Bootstrapping time series

Rationale: Twitter posts at irregular intervals, so mostly different samples for each time window (users, tweets, and sample size)!

- Bootstrapping our Twitter sample for each time period by randomly sampling with replacement
- Provides estimates of CIs to model uncertainty, in particular where N is low
U.S. County-level sentiment mapping tool
What do these numbers on the dots represent?
• Complete county-level sentiment data is now available in easy-to-parse JSON format as well as

• County-specific sentiment time series with null-model and 95% Confidence Intervals
Case studies for validation

We chose 3 locations hit by hurricanes

**validation**: do we find a sentiment signal at that time?

**Resilience**:
- Does sentiment return to baselines?
- How fast?

1. **Houston, TX**: ”Harvey”, Aug 17 – Sep 22, 2017
2. **Puerto Rico (entire territory)**: “Maria”, Sept 20, 2017
3. **Florida (entire state)**: “Irma”, Sept 10, 2017
FILE PHOTO: People stop on a highway near a mobile phone antenna tower (not pictured) to check for mobile phone signal, after the area was hit by Hurricane Maria, in Dorado, Puerto Rico September 23, 2017. REUTERS/Alvin Baez/File Photo
Comparing sentiment:
Entrepreneurs v. general population

Can we measure entrepreneurial sentiment via Twitter?
E-Ship validation efforts

1. Investigation of ability to detect sentiment specifically for online communities of entrepreneurs on Twitter

2. Cross-validation against state-level survey data of Subjective Well-Being in USA
Data:

Entrepreneur data

- 251 entrepreneur user accounts
- Each with up to 3,200 most recent tweets
- Tweet count distribution:

Null-model: Random tweets from one random day

- We chose 2015-03-23
- Over 36 million tweets

Null-Model: Random user timelines

- N=251
- Over 15,000 tweets
- Tweet count distribution:
Technique 1: VADER

Method

- Rate each “entrepreneur tweet” using VADER compound score
- Compare entrepreneur tweets to a day’s worth of random tweets
- Because we do see positive skew at the tweet level for entrepreneurs, a potential user-level score would be the VADER mean of all authored tweets

Right skew indicates higher levels of positive sentiment

Largely symmetric
Technique 2: Opinion Finder

Instead of evaluating all text sentiment, we rate “subjective”, i.e. personal statements only

Method

• OpinionFinder (OF) was developed to focus specifically on subjective sentiment

• For each entrepreneur timeline, we gather positively- and negatively-classified words by OF

• Entrepreneur individual sentiment score:
  • score = (#positive - #negative)/(#positive + #negative)
  • Collect only if number of classified words in the timeline is > 10

jbollon@Indiana.edu
Cross-validation Twitter vs. state-level SWB surveys:

Are we measuring Subjective Well-Being or another population sentiment via Twitter?
Behavioral Risk Factor Surveillance System

1. CDC - **nationwide telephone survey about health and risk behavior**
2. Started in 1984 with 15 states
3. Over 400,000 participants per year
4. Subjective well being - measure of quality of life - Confirmation (Owald et al.)
   a. Presence of positive emotion
   b. Absence of negative emotions
   c. Life satisfaction
   d. Fulfillment
   e. Positive functioning

**Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the U.S.A.**

jbollen@Indiana.edu
Subjective Well Being vs Sentiment

1. Do they correlate for each state? - hard to check because the later years are not complete
2. Used the mean of states over all years

no significant correlation at this point
Continuing work 2018-2019

- Refinement of sentiment analysis tool with objective to increase validity with respect to SWB and other social indicators
- Comparison of online sentiment data to more complete SWB data for cross-validation
- Applications of different sentiment analysis indicators: “text sentiment” vs. “personal SWB”
- Measuring regional and entrepreneurial resilience from longitudinal data
- Contributing validated sentiment data to E-Ship construction in collaboration with other partners