Detecting Sentiment and Tone in Chats

INRUG Spokane September, 2016

Why Sentiment Analysis?

- Sentiment Search (1)

- Social Media Mentions

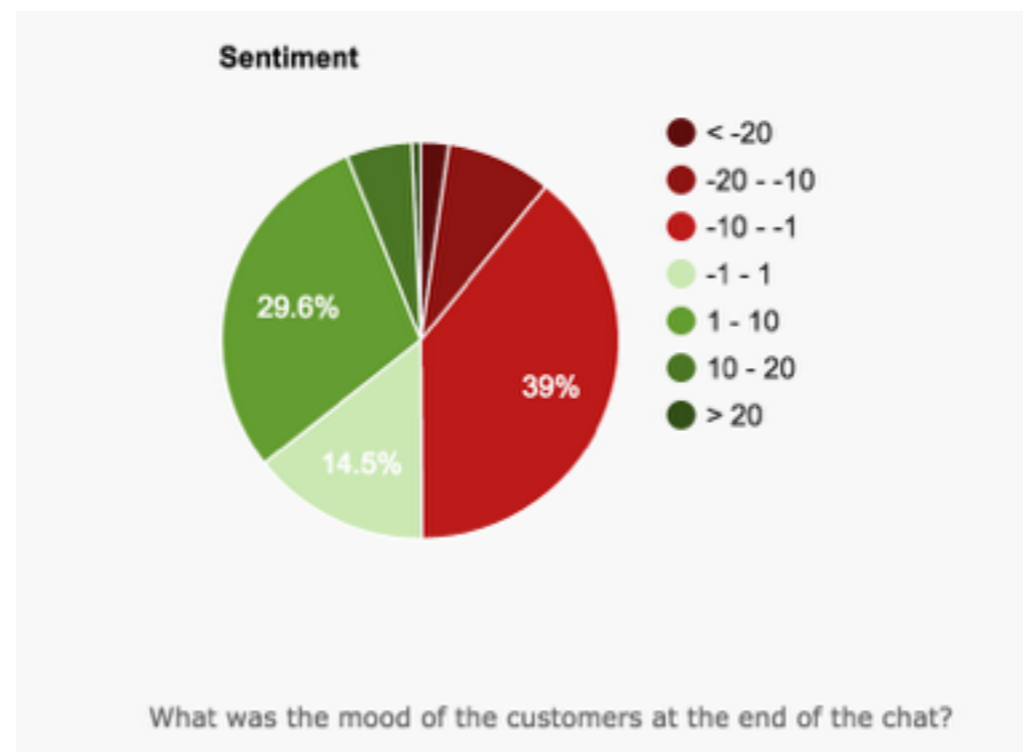- Customer Service Analytics

Sentiment Heuristics

- Look for words that indicate sentiment: "awesome", "lame", "helpful", "stupid"

- Look for phrasing: "It could have been awesome" "It didn't suck" "I'm not happy"

- Look for non-language indicators: "I AM NOT HAPPY"

Data-Driven Methods

- Use Machine Learning to Learn sentiment from data (1)

- Lots of text data around reviews, but not much labelled

- Humans agree about 80% on the sentiment of written reviews (2)

Conversational Sentiment

- If you have a conversation with someone, how do they feel at the end? Can you tell by what they say?

Sentiment

●	< -20
●	-20 - -10
●	-10 - -1
●	-1 - 1
●	1 - 10
●	10 - 20
●	> 20

29.6%

39%

14.5%

What was the mood of the customers at the end of the chat?

Sentiment Measures

- Discrete: positive, neutral, negative - Classification

- Continuous: [-1:1], - Regression and ranking methods. This is the most common measure.

  - -1 = completely negative,

  - 0 = neutral,

  - +1 = completely positive

Learning Sentiment

- RTextTools (3)

- Python NLTK (4)

- Keras Deep Learning (5)

Online Sentiment Tools

- Stanford Sentiment Analysis (6)

- Microsoft Text Analytics API (7)

- Alchemy API (IBM Watson) (8)

- Other tools: Python Textblob Sentiment Analysis

Sentiment and Tone

- Sentiment uses a score often called polarity. Very general.

- Tone indicates a more discrete emotional category.

- An emotional tone classifier for text was recently released by IBM Watson.

- Can we make a sentiment analyzer from tones?

Watson Tone Analyzer

- Five emotional Tones: Fear, Anger, Joy, Disgust, Sadness

- Scores are between [0.0:1.0]

- Document and sentence level tone information

- Information returned as a vector of values in JSON format.

Watson Tone Analyzer

- It's not clear what the tones actually mean. What's the difference between anger and disgust?

- There is no (that I could find) documentation on how the tones are trained or the algorithm used.

Tone Score

- Scores are returned as a JSON document

- Sentence: {"text": "Customer: Do you understand the problem I am having or is this a lost cause?"}

- {"tones": [{"score": 0.65351, "tone_id": "anger","tone_name": "Anger"},

  {"score": 0.214534,"tone_id": "disgust","tone_name": "Disgust"},

  {"score": 0.248598,"tone_id": "fear","tone_name": "Fear"},

  {"score": 0.028052,"tone_id": "joy","tone_name": "Joy"},

  {"score": 0.423949,"tone_id": "sadness","tone_name": "Sadness"}

  ],"category_id": "emotion_tone","category_name": "Emotion Tone"}

Data

- I used 6000 chats between users and an automated agent. polarity was from the python textBlob sentiment analysis.

```
tone = read.csv("../outtest.csv", strip.white=TRUE)

df1 <- data.frame("Text" = tone$Text,

                  "Anger" = tone$Anger,

                  "Disgust" = tone$Disgust,

                  "Fear" = tone$Fear,

                  "Joy" = tone$Joy,

                  "Sadness" = tone$Sadness,

                  "polarity" = tone$polarity,

                  "subjectivity" = tone$subjectivity)
```

Data

1 I'm not understanding that the audio need to be connected separately fr "Video between DVI-D and HDMI ports is compatible with a physical adapter, but audio would need to be connected separately."

2 How can I help you with your equipment?

3   On this displayed page you will see information about connecting your Digital Receiver using various types of connections, including HDMI, Component, S-Video and others.

4 question, can't get audio to work on my tv after hooking up the hdmi high speed ethernet to tv and cable box what am i doing wrong?

5 chat rep

6 Did you know that I?m able to quickly help with most issues without waiting for a live person?

 Give me a try by asking a brief question like the examples below:

7 How do I get to emails

8 Your email is displayed on the page I?ve opened. You can always access your email by selecting the envelope icon in the top right corner of the page after signing in.

 If you need help setting up a 3rd party mail client like Outlook or Mac Mail or accessing your mail through a mobile device, I can help.

9 insure computer is protected

10 understand that you are interested in information about a computer. Please click one of the links below to learn more.

Data

	Anger	Disgust	Fear	Joy	Sadness	polarity	subjectivity
1	0.305193	0.476463	0.182303	0.321236	0.318414	0.0000000	0.1428571
2	0.342848	0.350850	0.159783	0.198855	0.159902	0.0000000	0.0000000
3	0.361415	0.436878	0.387802	0.074279	0.174068	0.0000000	0.2500000
4	0.480038	0.350837	0.152525	0.185456	0.231341	-0.1700000	0.7200000
5	0.274187	0.154859	0.158057	0.370935	0.270699	0.0000000	0.0000000
6	0.344532	0.203722	0.183477	0.144933	0.504681	0.2939394	0.4916667
7	0.388716	0.183042	0.174494	0.149420	0.280324	0.0000000	0.0000000
8	0.331688	0.235099	0.200010	0.359153	0.126253	0.2619048	0.3452381
9	0.299381	0.376371	0.256596	0.175973	0.134950	0.0000000	0.0000000
10	0.567752	0.440215	0.123935	0.091082	0.200724	0.3750000	0.5000000

Analysis

- Created 5 models: one linear model for each tone.

      model1 <- lm(polarity~Anger, data=df1)

- Checked the summary values for each model, e.g,

> summary(model1)

Call:

lm(formula = polarity ~ Anger, data = df1)

Residuals:

    Min   1Q   Median   3Q   Max

-1.09186 -0.08963 -0.08175 0.04444 0.92277

Coefficients:

          Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.100094   0.007162 13.975   <2e-16 ***

Anger   -0.040007   0.018601 -2.151   0.0315 *

Residual standard error: 0.1988 on 6012 degrees of freedom

Multiple R-squared: 0.0007688,   Adjusted R-squared: 0.0006026

F-statistic: 4.626 on 1 and 6012 DF, p-value: 0.03153

Analysis

- Ran Anova for all 5 models:

```
> anova(model1, model2, model3, model4, model5)

Model 1: polarity ~ Anger

Model 2: polarity ~ Joy

Model 3: polarity ~ Disgust

Model 4: polarity ~ Sadness

Model 5: polarity ~ Fear

  Res.Df     RSS Df Sum of Sq F Pr(>F)

1   6012 237.59

2   6012 230.65  0     6.9323

3   6012 237.55  0    -6.8947

4   6012 233.82  0     3.7233

5   6012 235.02  0    -1.1985
```

Analysis

- Each model has a very low Standard and Adjusted R-squared indicating that the model does not explain variance.

- The Tones individually do not predict the sentiment in this case.

Combining Tones

- Created one more model with all 5 tones. Also fails to predict sentiment.

```
model6 <- lm(polarity~Anger+Joy+Disgust+Sadness+Fear, data=df1)

> summary(model6)

(Intercept) -0.01728    0.02934  -0.589 0.555887

Anger        0.10623    0.02897   3.667 0.000248 ***

Joy          0.31424    0.03117  10.081  < 2e-16 ***

Disgust      0.10269    0.02237   4.591 4.51e-06 ***

Sadness     -0.05762    0.02738  -2.104 0.035396 *

Fear        -0.08161    0.03012  -2.709 0.006767 **

Residual standard error: 0.1939 on 6008 degrees of freedom

Multiple R-squared:  0.05007, Adjusted R-squared:  0.04928

F-statistic: 63.34 on 5 and 6008 DF,  p-value: < 2.2e-16
```

Other Experiments

- Are there other factors that, added to tone, could improve the predictions?

- Try some different sentiment analyzers and see if these results are consistent.

- What do these tones mean? In the extremes of the tone scores they seem to make sense.

References

1. Pang & Lee, http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf

2. Agreement study, http://people.cs.pitt.edu/~wiebe/pubs/papers/emnlp05polarity.pdf

3. R Tutorial, https://www.r-bloggers.com/sentiment-analysis-with-machine-learning-in-r/

4. NLTK, http://text-processing.com/demo/sentiment/

References (Cont.)

- 5. Keras tutorials, https://github.com/fchollet/keras/blob/master/examples/imdb_lstm.py

- 6. Stanford SA, http://nlp.stanford.edu:8080/sentiment/rntnDemo.html

- 7. Microsoft Text Analytics, https://www.microsoft.com/cognitive-services/en-us/text-analytics-api

- 8. Alchemy API, http://www.alchemyapi.com/products/alchemylanguage/sentiment-analysis

- 9. Watson Tone Analyzer, https://www.ibm.com/watson/developercloud/tone-analyzer.html