Cross-lingual SA

Lyle Ungar

University of Pennsylvania

Translating sentiment is subtle

- **Machine translation is not terrible**
  - But cultures us different words
    - Hahahahaha
  - And sometimes have different feelings about the same words
    - España

Does ‘well-being’ translate on Twitter?

English-Spanish translation

- **Label English or Spanish Tweets for**
  - Positive emotion
  - Engagement
  - Relationships
  - Meaning
  - Accomplishment
- **Build models**
  - Translate them and **compare**

Some words evoke different feelings

Meaning

Not accomplishment

Not meaning

Not relationship

Not positive emotions

PERMA	term	weight (en)	weight (es)	% chg	
POS_M (en)	mundo* (world)	0.42	-0.18	143	
NEG_A (en)	odio** (hate)	0.29	2.19	87	
NEG_M (en)	nadie*** (no one)	0.23	0.24	4.2	No change
NEG_R (es)	sad** (triste)	1.70	0.0012	100	
NEG_P (es)	hate*** (odio)	1.81	1.75	3.3	No change

English-Spanish translation

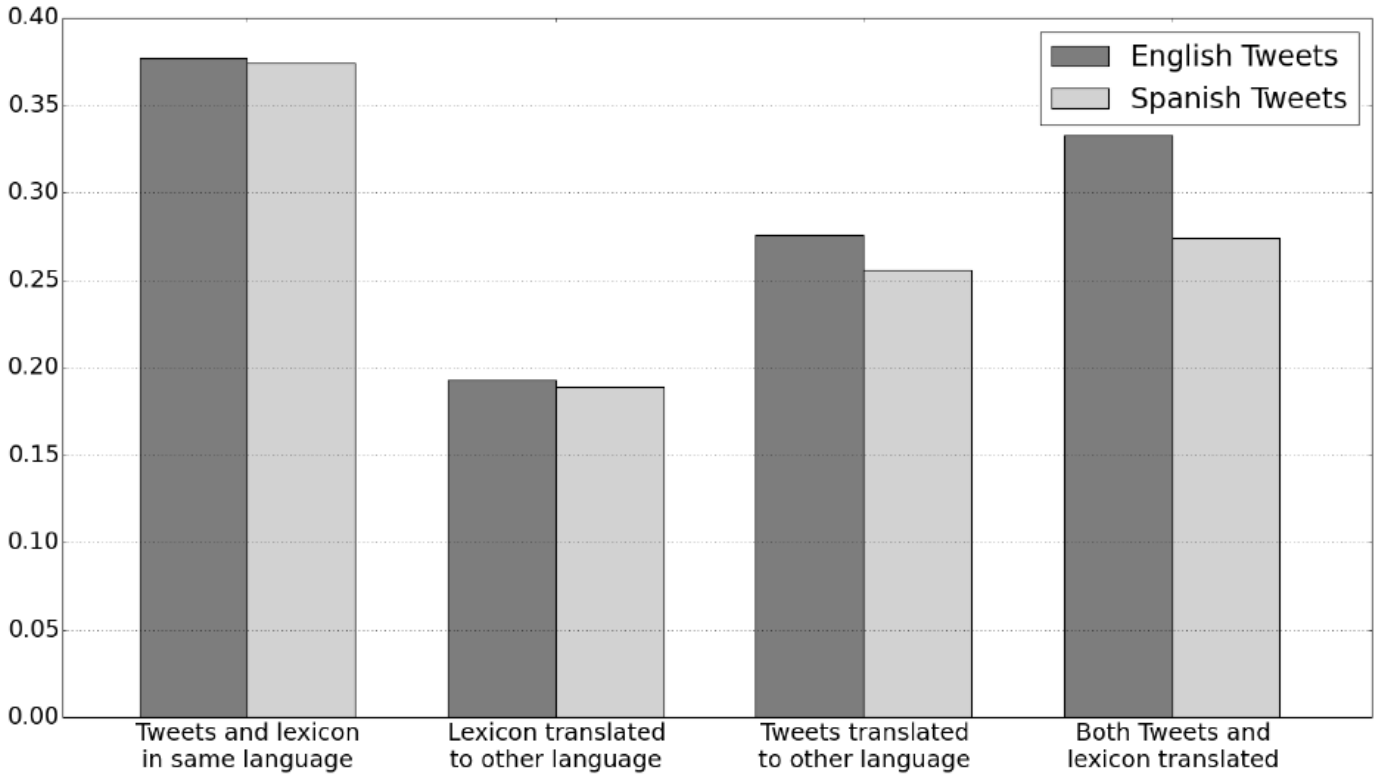
- **The biggest problem is 'missing terms'**

source lang	correct trans	missing terms	opp sign	weight diff
English	83%	81%	0.5%	6.9%
Spanish	74%	91%	0.0%	4.8%

Translating sentiment is subtle

- **Cultures are different**
  - La selectividad — SAT
  - Cinco de Mayo — 4th of July
  - Iker Casillas — Michael Jordan
  - Gilipollas, tonto (douchebag, fool) — kill, stupid — negative emotion
  - Hermana (sister) — friend — relationships

Best: train in your target language



Twitter is not mono-lingual

- **English speaking countries differ**
  - behaviour, theatre, mum, boot
  - 4th of July, GRE, SAT
  - **bloody, brilliant**
- **Different products differ**
  - Blue, small
- **English posts often aren't pure English**
  - "Low IQ" Facebook users in California
  - Extreme personality predictions in some states

Domain Adaptation

- **Structured correspondence learning**
  - Find words associated with *pivots* in both domains
    - **Pivots:** excellent, awful, …
    - **Associated words:**

domain\polarity	negative	positive
books	plot <num>_pages predictable reading_this page_<num>	reader grisham engaging must_read fascinating
kitchen	the_plastic poorly_designed leaking awkward_to defective	excellent_product espresso are_perfect years_now a_breeze

Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification

Domain Adaptation

- **Structured correspondence learning**
  - Find words associated with *pivots* in both domains
  - Θ maps original words to ones in new domain

$$\min_{\mathbf{w},\mathbf{v}} \sum_i L\left(\mathbf{w}'\mathbf{x}_i + \mathbf{v}'\theta\mathbf{x}_i, y_i\right) + \lambda||\mathbf{w}||^2 + \mu||\mathbf{v}||^2 \ ;$$

w weights on original features
v weights on projected features – *adjust on target domain*
$\mathbf{y}_i$ labels

Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification

Multimodal SA

Lyle Ungar

University of Pennsylania

Images reflect sentiment

Images reflect personality

(a) Extraverted.

(b) Conscientious.

Figure 1: Example Twitter profile pictures for users scoring high in a personality trait.

Image features

Contrast	
Saturation	High indicates vividness and chromatic purity – more appealing to the human eye
Sharpness	Measures coarseness or the degree of detail contained in an image, a proxy for the quality of the photographing gear
Blur	Low blur for higher quality images
Grayscale	If the image is in grayscale – Black/White photos are more artistic
Naturalness	The degree of correspondence between images and human perception
Brightness	
Colorfulness	The difference against gray
Color Emotions	Affective tone of colors, represented by 17 color histogram features
RGB Colors	
Hue	

Deep nets detect face sentiment

- Joy, surprise, anger, disgust, fear sadness

imagga.com

clairifai.com

https://www.kairos.com/emotion-analysis-api

https://www.microsoft.com/cognitive-services/en-us/emotion-api

Face Detection

Find and track faces in any video, photo or image.

Face Identification

Search for face matches. Answers: “Who is this?”.

Face Verification

Search for someone. Answers: “Is this Elizabeth?”.

Emotion Detection

Joy, anger, disgust, sadness, fear and surprise.

Age Detection

Detects age groups; child, young-adult, adult, or senior.

Gender Detection

Detects gender of each face found; female or male.

Multi-face Detection

Detects individuals, crowds, audiences and groups.

Attention Measurement

Total attention time, glances, blinks, and attention span.

Facial Features

Detects eyes, eyebrows, nose, mouth, and many more.

Sentiment Detection

Understand positive, negative and

Face Grouping

Group faces for easy searching,

Ethnicity Detection

Understand the diversity of the

https://
www.kairos.com

Videos reflect sentiment / emotion

- **Videos (and audio) reveal emotion**
  - http://www.affectiva.com
- **Increasingly done with neural nets**
  - Recurrent Neural Networks for Emotion Recognition in Video

Model	Training	Validation
Activity	0.983	0.266
Audio	0.418	0.332
Aggregated CNN	0.505	0.350
RNN	0.848	**0.396**

Emotion

- **Phone-based apps**

AffdexMe

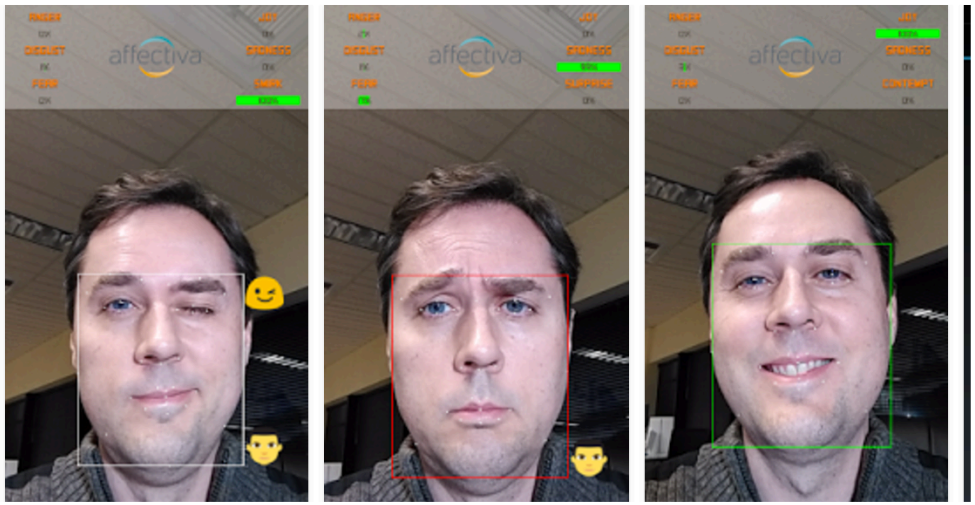
Affectiva Inc. Libraries & Demo

★★★★★ 122

Everyone

This app is compatible with your device.

Installed



AffdexMe demonstrates the use of Affectiva's Affdex Software Developer Kit (SDK), which lets developers emotion-enable their apps and digital experiences.

Conclusions

- **Be very careful about '*translating*' sentiment**
- **Sentiment in images**
  - Content extraction often challenging
    - But color spectra easy to pull
  - Pre-trained deep networks give useful features
    - Faces: smiling, bored, …
    - Objects: target of the sentiment and how it is viewed
- **Sentiment in video – strong but tricky to get**

Wrap-up

Lyle Ungar, University of Pennsylania

Tutorial outline

- **Core sentiment analysis (SA) methods**
  - **Simple:** using lexica (dictionaries)
  - **Aspect-based:** using information extraction or parsing
- **Machine learning for SA**
  - Unsupervised: open language SA – **DLA & LDA**
  - Supervised: regression and deep learning
- **SA extensions**
  - Post, person and community: **emotion, personality**
  - Multilingual, multimodal

Thank you!!!!

Questions?

ungar@cis.upenn.edu