Emotion Classification

Keshav Joshi

Emotion Classification: Background

Background

- Many emotion classes:
  - Plutchnik's 8
  - Valence-Arousal-Dominance
  - Ekman's 6
  - Big 5 (We chose this one!)
  - and more … wiki

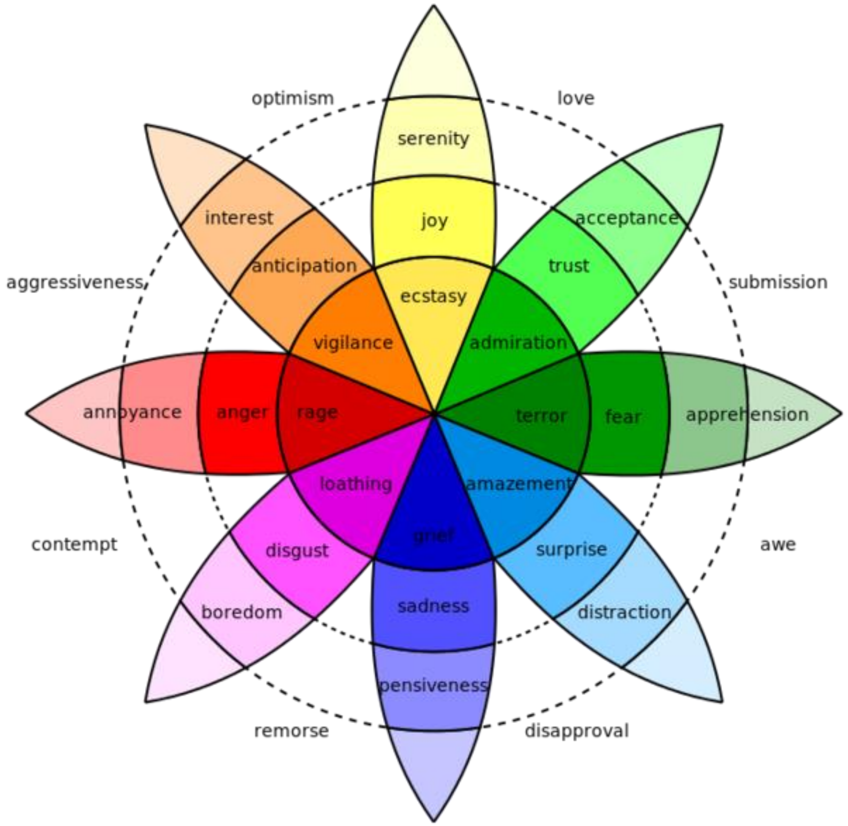
Plutchnik

courtesy: https://www.6seconds.org/wp-content/uploads/2018/04/Screen-Shot-2018-07-02-at-10.55.00-AM.png

Plutchnik

courtesy: https://www.6seconds.org/wp-content/uploads/2018/04/Screen-Shot-2018-07-02-at-10.55.00-AM.png

Added dimension of intensity / arousal



courtesy: https://upload.wikimedia.org/wikipedia/commons/c/ce/Plutchik-wheel.svg

Valence Arousal Dominance

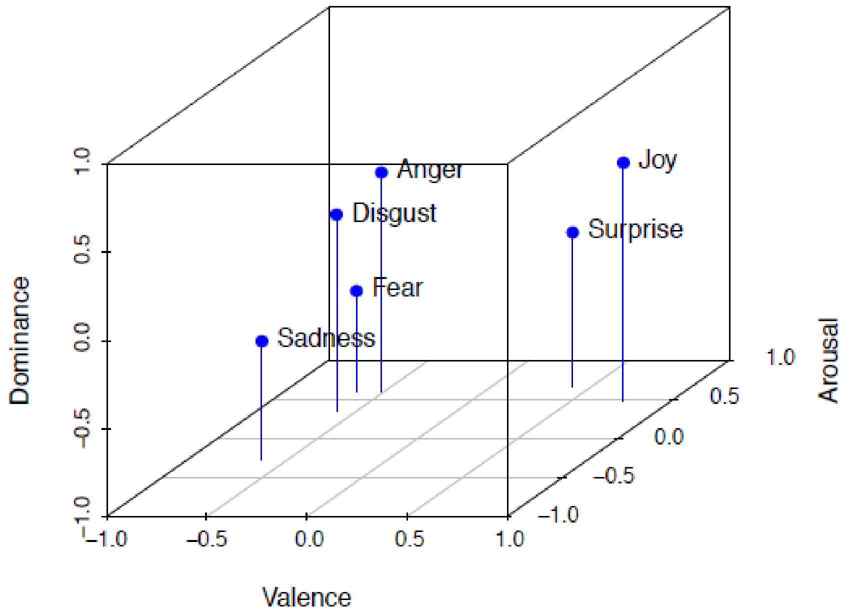


Figure 1: The affective space spanned by the three VAD dimensions. As an example, we here include the positions of Ekman's six Basic Emotions as determined by Russell and Mehrabian (1977).

- Valence: positive or negative emotion
- Arousal: intensity of emotion
- Dominance: how assertive the speaker is

Courtesy: Buechel, Sven, and Udo Hahn. "EMOBANK: Studying the impact of annotation perspective and representation format on dimensional emotion analysis." *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Vol. 2. 2017.

Across emotion models

Plutchnik	Ekman	Big-5
Anger	Anger	Anger
Fear	Fear	Fear
Disgust	Disgust	Disgust
Sadness	Sadness	Sadness
Joy	Joy	Joy
Surprise	Surprise	
Anticipation		
Trust		

Emotion Classification: Machine Learning

Emotion Prediction Pipeline

- Gather Data
- Process Data
- Train Model
- Deploy Model: API

Datasets

Dataset	Granularity	Annotation	Size	Topic	Source	Avail.
AffectiveText	headlines	E + V	1,250	news	Strapparava (2007)	D-U
Blogs	sentences	E + ne + me	5,025	blogs	Aman (2007)	R
CrowdFlower	tweets	E + CF	40,000	general	Crowdflower (2016)	D-U
DailyDialogs	dialogues	E	13,118	multiple	Li et al. (2017)	D-RO
Electoral-Tweets	tweets	P	4,058	elections	Mohammad (2015)	D-RO
EmoBank	sentences	V+A+D	10,548	multiple	Buechel (2017a)	CC-by4
EmoInt	tweets	E − DS	7,097	general	Mohammad (2017b)	D-RO
Emotion-Stimulus	sentences	E + shame	2,414	general	Ghazi et al. (2015)	D-U
fb-valence-arousal	faceb. posts	V+A	2,895	questionnaire	Preoţiuc (2016)	D-U
Grounded-Emotions	tweets	HS	2,585	weather/events	Liu et al. (2017)	D-U
ISEAR	descriptions	E + SG	7,665	events	Scherer (1994)	GPLv3
Tales	sentences	E	15,302	fairytales	Alm et al. (2005)	GPLv3
SSEC	tweets	P	4,868	general	Schuff et al. (2017)	D-RO
TEC	tweets	E ±S	21,051	general	Mohammad (2012)	D-RO

Table 1: Selection of resources for emotions analysis. Ann. refers to the following annotation schemata: [E] *Ekman: anger, disgust, fear, joy, sadness, surprise*, [P] *Plutchik: anger, disgust, fear, joy, sadness, surprise, trust, anticipation*, [CF] *enthusiasm, fun, hate, neutral, love, boredom, relief, empty*, [DS] *disgust, surprise*, [JS] happy, sad, [V] *valence*, [A] *arousal*, [D] *dominance*, [SG] *shame, guilt*, [±S] *positive surprise, negative surprise*, [ne] *no emotion* [me] *mixed emotion* and Availability refers to the following [D-RO] *Available to download, research only*, [D-U] *Available to download, unknown licensing*, [R] *Available upon request*, [GPLv3] *GNU Public License version 3*, [CC-by 4] *Creative Commons Attribution version 4.0*

Courtesy: Bostan, Laura Ana Maria, and Roman Klinger. "An Analysis of Annotated Corpora for Emotion Classification in Text." *Proceedings of the 27th International Conference on Computational Linguistics*. 2018.

Process Data

- Map emotion labels
- Tokenize text (nltk.tokenize.TweetTokenizer)
  - [‘I like learning’] => [‘I’, ‘like’, ‘learning’]
- Stem text (nltk.stem.snowball.PorterStemmer)
  - ‘learning’ => ‘learn’
- Remove special characters, multiple consecutive digits
- TfidfVectorizer: replace word with a bag of words matrix of word-frequencies scaled by TFIDF score
  - to minimize the effect of words like ‘the’ which are very frequent
  - Note: stopwords are not removed

Process Data: map emotion labels

Emotions	<= Labels (mapped to emotions)
joy	enthusiasm, love, surprise, relief, fun, happiness
sadness	sadness
anger	hate
fear	worry
disgust	disgust
neutral	empty, boredom
dropped	shame, guilt (ISEAR)

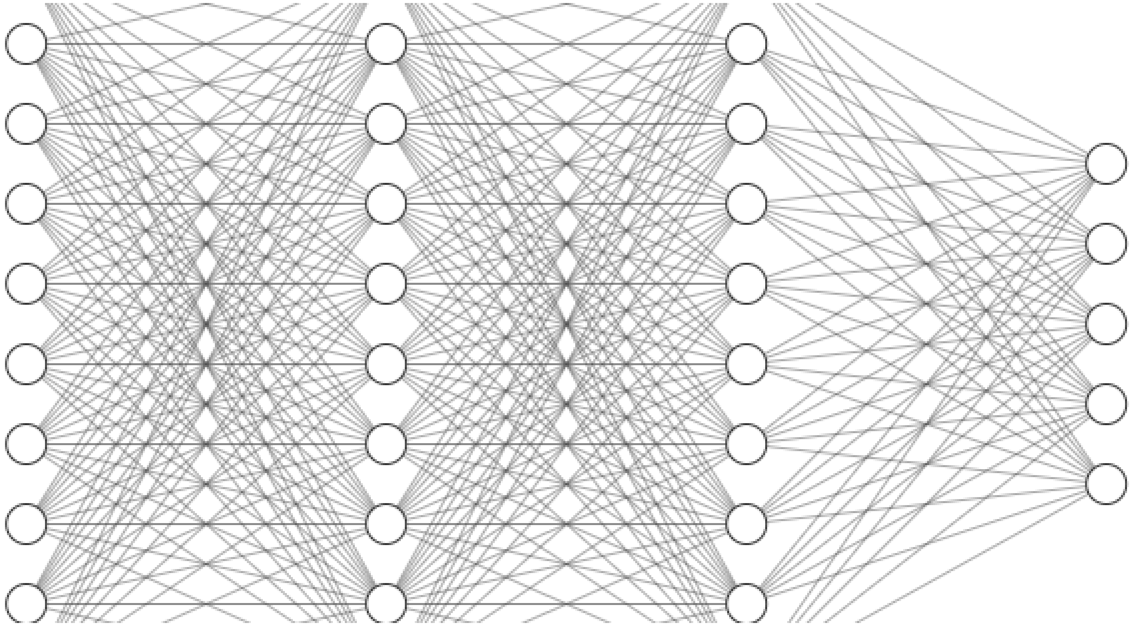
Emotion mappings from each dataset

tweets		ISEAR	
happiness	15299	joy	1094
neutral	9644		
fear	8459	fear	1095
sadness	5165	sadness	1096
anger	1433	anger	1096
		disgust	1096

Emotion frequencies from each dataset

Train model: architecture

Preprocessed text ➔ 100 x 100 x 100 Neural Network



➔ Multi-class (5) softmax probability output

Train Model: results (bag-of-words)

	precision	recall	pct	mse
anger	0.31	0.35	0.43	0.96
disgust	0.33	0.51	0.43	0.88
fear	0.34	0.39	0.43	1.10
joy	0.53	0.71	0.43	0.75
sadness	0.28	0.36	0.43	1.07

Note: F1 scores are not reported here, but aggregate ones for this model are later on
Gaussian Naïve-Bayes and Random Forest classfiers were also tested, but to worse results

Train Model: results (fastText sentence-embedding)

	precision	recall	f1	pct	mse
anger	0.11	0.16	0.16	0.32	1.57
disgust	0.23	0.28	0.28	0.32	1.47
fear	0.25	0.31	0.31	0.32	1.39
joy	0.43	0.54	0.54	0.32	1.08
sadness	0.19	0.26	0.26	0.32	1.46

Basic LSTM/CNN architecture on word-level fastText embeddings was also tested, but the scores have not been tabulated

Train model: human accuracy

- Human accuracy:

F1-score	Anger	Disgust	Fear	Joy	All
	46%	21%	51%	58%	45%

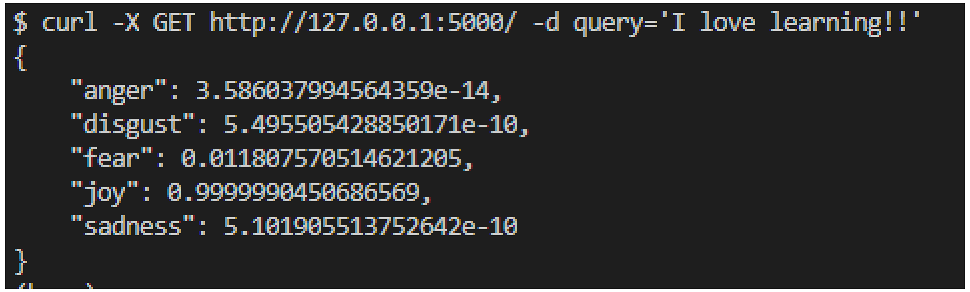
Courtesy: Klinger, Roman. "Orphée de Clercq, Saif M. Mohammad, and Alexandra Balahur. 2018. Iest: Wassa-2018 implicit emotions shared task." *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Brussels, Belgium. Association for Computational Linguistics*.

- Overall model results:

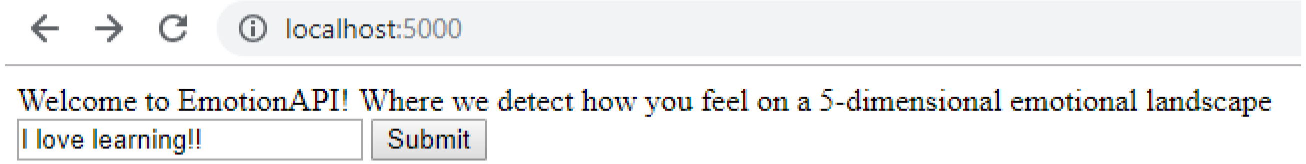
	Bag-of-words NN	Embedding NN	Human benchmark
F1-score	41%	32%	45%

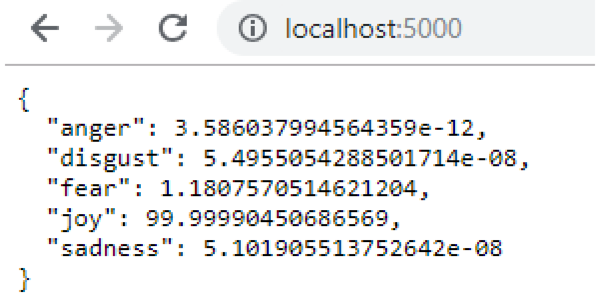
API (runs locally)

- Command-line API:
    - GET Request: curl -X GET http://127.0.0.1:5000/ -d query='I love learning!'
    - Results:



- Web API:
    - Results:

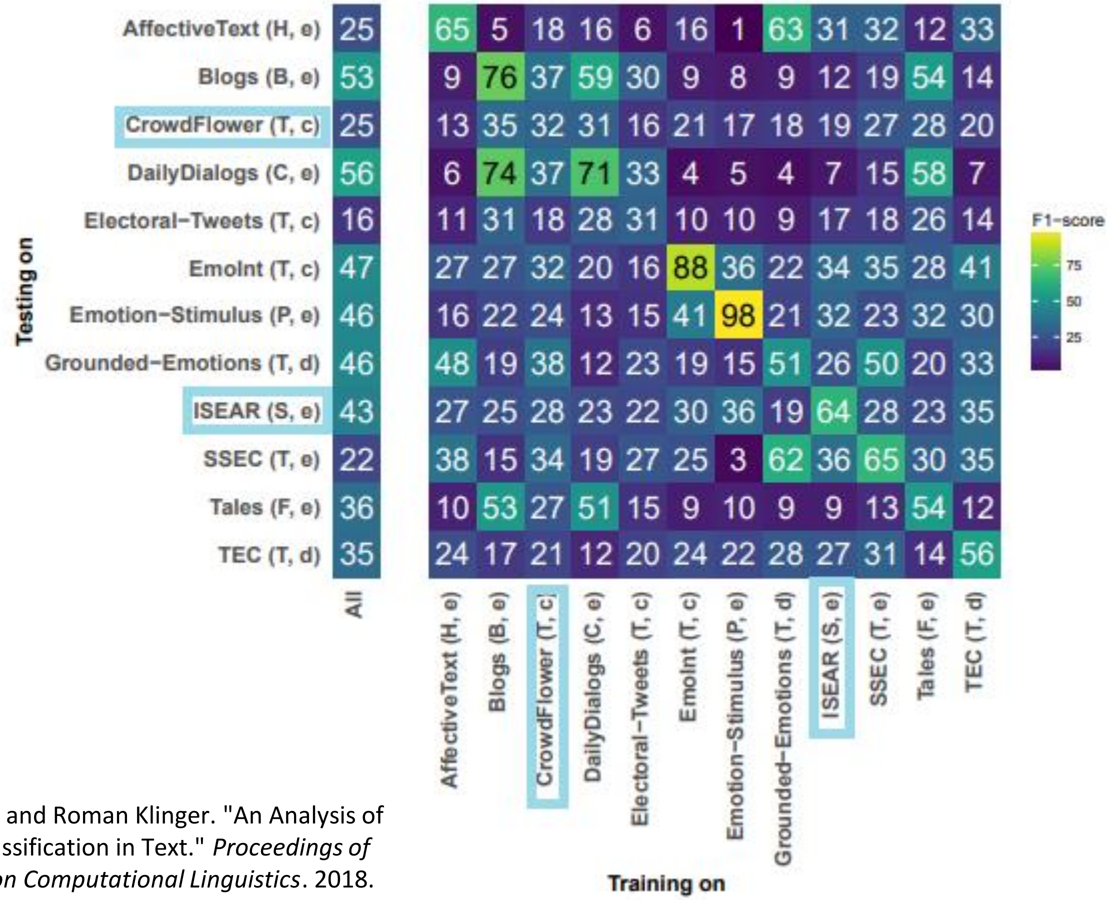




Emotion Classification: State-of-the-art

Text => emotion is just a hard problem?

- Results from training a logistic regression model on one dataset and testing on another
- cross-corpus high scores are rare



Courtesy: Bostan, Laura Ana Maria, and Roman Klinger. "An Analysis of Annotated Corpora for Emotion Classification in Text." *Proceedings of the 27th International Conference on Computational Linguistics*. 2018.

Text => emotion is just a hard problem?

- Results from classifying corpus of ~3K tweets across many research teams
- same corpus w/ a human benchmark reported earlier
- It is possible to get decent scores, ones that far surpass human benchmarks

	Joy			Sadness			Disgust			Anger			Surprise			Fear		
Team	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1	P	R	F_1
Amobee	**82**	**82**	**82**	70	**68**	**69**	**73**	70	**72**	62	**66**	**64**	**66**	70	68	**77**	73	**75**
IIIDYT	79	81	80	**71**	67	69	70	**71**	71	**66**	63	64	**66**	**71**	**68**	76	**74**	75
NTUA-SLP	81	77	79	71	66	69	72	70	71	63	64	63	62	71	67	75	73	74

Courtesy: Klinger, Roman. "Orphée de Clercq, Saif M. Mohammad, and Alexandra Balahur. 2018. Iest: Wassa-2018 implicit emotions shared task." *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Brussels, Belgium. Association for Computational Linguistics*.

Third-party Emotion APIs

- IBM Watson
  - Includes sentiment and big-5
- LIWC (word-emotion hash table)| Receptiviti (API service)
  - Includes sentiment and Analytical Thought (CI?)
- Amazon Comprehend
  - Includes sentiment and big-5
- Qemotion API
  - Includes sentiment, VAD, and Ekman's 6
- WordNet Affect (multiple emotion labels)
  - https://stackoverflow.com/questions/27943396/using-wn-affect-to-detect-emotion-mood-of-a-string/27945838#27945838
- How are all these models built and validated? proprietary information…