Emotion Recognition in Images and Text

Agata Lapedriza

alapedriza@uoc.edu / agata@mit.edu



Associate Professor









Visiting Researcher



Recognizing others emotions... Why is this capacity useful?

Recognizing others emotions...

Why is this capacity useful?

- Social Interactions
- Detecting people's needs
- Predicting people's reactions



Emotions



Cognition

https://www.maxpixel.net/Calm-Smiley-Ball-Angry-Anxiety-Emoticon-Anger-2979107 https://www.maxpixel.net/Quiz-Think-Question-Thinking-Brain-Answer-2004314

Al in Science Fiction Movies









Emotion Recognition

A lot of signals in our bodies change when our emotions change



Emotion Recognition



Emotion Recognition



(ex: cameras, ...)

Facial expression analysis



AffdexMe app :) Affectiva https://www.affectiva.com





Smile (0.99), Attention (0.54)



Facial expression analysis

1. Non-frontal Faces and Partial Occlusions.



Surprise Mouth Open (1.00)





Facial expression analysis

- 1. Non-frontal Faces and Partial Occlusions.
- 2. Some Facial Expressions are not related to our emotions, but to our actions.
- 3. Difficulty in giving Emotional Meaning to an isolated Facial Expression.



Aviezer, H., Hassin, R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. Angry, disgusted or afraid? Studies on the malleability of emotion perception. Psychological Science, 19, 724-732 (2008a)



Aviezer, H., Hassin, R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. Angry, disgusted or afraid? Studies on the malleability of emotion perception. Psychological Science, 19, 724-732 (2008a)













Apparent Emotional States Recognition



Apparent Emotional States Recognition





Apparent Emotional States Recognition





Confidence feeling of being certain; conviction that an outcome will be favorable; encouraged; proud

From pictures to emotions





Deep Learning Model Emotions

From pictures to emotions





Challenge: Training Data

Collecting images



• • •

- Images manually downloaded from search engines
- Images from other public datasets: MSCOCO, Ade20k

Back (mage 1 of 20)	<pre>mace (well being and relaxed/no worry/positive sensation/satisfied) fection (fond feelings/tenderness/love/compassion) pectation (state of anticipating/hoping on something or someone) teem (favorable opinion or judgment/gratefulness/admiration/respect) onfidence (feeling of being certain/proud/encouraged/optimistic) rgagement (occupied/absorbed/interested/paying attention to something) easure (feeling of delight in the senses) appiness (feeling delighted/enjoyment/amusement) citement (pleasant and excited state/stimulated/energetic/enthusiastic) urprise (sudden discovery of something unexpected) offering (distressed/perturbed/anguished) sapproval (think that something is wrong or reprehensible/contempt/hostile) arning (strong desire to have something/jealous/envious) tigue (weariness/tiredness/sleepy) in (physical suffering) oubt/Confusion (difficulty to understand or decide/sceptical/lost) ar (feeling afraid of danger/evil/pain/horror) unerability (feeling of being physically or emotionally wounded) squitement (unpleasant restlessness/tense/worried/upset/stressed) moyance (bothered/iritated/impatient/troubled/frustrated) neger (intense displeasure or rage/furious/resentful) sgust (feeling dislike or repulsion/feeling hateful) dness (feeling unhappy/grief/disappointed/discouraged) sconnection (not participating/indifferent/bored/distracted) nbarrassment (feeling ashamed or guilty)</pre>
---------------------	--

	Valence: Negative vs. Positive
	Negative (unpleasant)
	Arousal (awakeness): Calm vs. Ready to act
	Calm
	Dominance : <i>Dominated</i> vs. <i>In control</i>
	Dominated (no control) 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
	Gender and age of the person in the yellow box
And a second sec	
	O Male O Female
	O Male O Female O Kid (0-12) O Teenager (13-20) O Adult (more than 20)







Anticipation
Excitement
Engagement





Yearning

Pleasure









23,571 Annotated Images 34,320 Annotated People



Deep Learning Model





AnticipationEEngagementC

Excitement Confidence





Pleasure Happiness Affection


Happiness







Pleasure Affection Happiness

Pleasure Disaproval Doubt/Confusion Disquietment Surprise Sensitivity Aversion Fatigue Sadness Esteem



R. Kosti, J.M Alvarez, A. Recasens, A.Lapedriza. "Context based Emotion Recognition using EMOTIC dataset". IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2019.

R. Kosti, J.M Alvarez, A. Recasens, A.Lapedriza. "Emotion Recognition in Context". Computer Vision and Pattern Recognition (CVPR), 2017.

Text Sentiment Analysis

The Red House Restaurant

98 Winthrop St, Cambridge, MA, Estats Units d'Amèrica

4,2 ★★★★ 391 comentaris @



Frederick Wright

Local Guide · 957 comentaris · 259 fotos

★★★★★ fa 2 setmanes

Have walked by this charming historic house on Winthrop Street for many years, but never had a chance to stop in until recently. What a delight! Terrific menu, tight but with options for everyone, great craft cocktails, unpretentious atmosphere of mostly locals and college professors chatting over dinner. We sat out on the porch where the acoustic were a bit milder and had a flawless mid-week meal. The small 'shared' plates are a bit smaller than expected, so I wish I'd ordered one or two more, but the main courses are generous.



Text Sentiment Analysis

What a delight! Terrific menu, great craft cocktails, unpretentious atmosphere of mostly locals and college professors chatting over dinner.





Text Sentiment Analysis

Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm

Bjarke Felbo¹, Alan Mislove², Anders Søgaard³, Iyad Rahwan¹, Sune Lehmann⁴

¹Media Lab, Massachusetts Institute of Technology ³College of Computer and Information Science, Northeastern University ³Department of Computer Science, University of Copenhagen ⁴DTU Compute, Technical University of Denmark

Abstract

NLP tasks are often limited by scarcity of manually annotated data. In social media sentiment analysis and related tasks, researchers have therefore used binarized emoticons and specific hashtags as forms of distant supervision. Our paper shows that by extending the distant supervision to a more diverse set of noisy labels, the models can learn richer representations. Through emoji prediction on a dataset of 1246 million tweets containing one of 64 common emojis we obtain state-of-theart performance on 8 benchmark datasets within sentiment, emotion and sarcasm detection using a single pretrained model. Our analyses confirm that the diversity of our emotional labels yield a performance improvement over previous distant supervision approaches.

1 Introduction

A variety of NLP tasks are limited by scarcity of manually annotated data. Therefore, co-occurring emotional expressions have been used for distant supervision in social media sentiment analvsis and related tasks to make the models learn useful text representations before modeling these tasks directly. For instance, the state-of-the-art approaches within sentiment analysis of social media data use positive/negative emoticons for training their models (Deriu et al., 2016; Tang et al., 2014). Similarly, hashtags such as #anger, #joy, #happytweet, #ugh, #yuck and #fml have in previous research been mapped into emotional categories for emotion analysis (Mohammad, 2012). Distant supervision on noisy labels often enables a model to obtain better performance on the target task. In this paper, we show that extend-

ing the distant supervision to a more diverse set of noisy labels enables the models to learn richer representations of emotional content in text, thereby obtaining better performance on benchmarks for detecting sentiment, emotions and surcasm. We show that the learned representation of a single pretrained model generalizes across 5 domains.

Table 1: Example sentences scored by our model. For each text the top five most likely emojis are shown with the model's probability estimates. -----I leve mom's cooking I leve how you never reply back. I leve cruising with my homies · · · · · · Have messing with yo mindit I love you and now you're just gone. 🤨 🤗 😝 😫 😣 This is shit A 11 d . . This is the shit

Emojojs are not always a direct labeling of emotional content. For instance, a positive emoji may serve to disambiguate an ambiguous sentence or to complement an oltervise relatively negative text. Kunnemna et al. (2014) discuss a similar duality in the use of emotional hashtags such as #nice and *Hame*. Nevertheless, our work shows that emojis can be used to classify the emotional content of texts accurately in many cases. For instance, our DeepMoji model captures varied usages of the word 'love' as well as slang such as 'this is the shit' being a positive statement (see Table 1). We provide an online demo at deepmoji.mit.edu to allow others to explore the predictions of our model.

Contributions We show how millions of readily available emoji occurrences on Twitter can be used to pretrain models to learn a richer emotional

DeepMoji

mit media lab machines

Felbo, B., Mislove, A., Sogaard, A., Rahwan, I. and Lehmann, S., 2017. Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. *arXiv preprint arXiv: 1708.00524.*

DeepMoji



1246 million tweets containing, at least, one of the 64 common emojis

Example: U2 Wonderful concert yesterday in Barcelona 😍



😬 DeepMoji Artificial emotional intelligence

DeepMoji has learned to understand emotions and sarcasm based on millions of emojis. Here's a video explaining a bit more. Type a sentence to see what our AI algorithm thinks.



Interactive scenarios

The capacity of recognizing emotions in sentences is also interesting in interactive scenarios.



Goal Oriented vs. Open Domain

Interactive scenarios



Open Domain

Open Domain Dialog Systems

How do we evaluate open domain dialog systems?

OPTION 1: Automatic text metrics (word overlap metrics; ex: BLEU score), Embedding-distance based metrics (ex: Average, Greedy, Extrema)

Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. How not to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2122–2132, 2016.

Open Domain Dialog Systems

How do we evaluate open domain dialog systems?

OPTION 1: Automatic text metrics (word overlap metrics; ex: BLEU score), Embedding-distance based metrics (ex: Average, Greedy, Extrema)

OPTION 2: Human Evaluation

Human evaluation of open domain dialog systems

The common practice is to use 1-turn evaluation

Context:

I think I want an ipad2. but I already have an ibook. Want vs. need.

Response:



A human rates how good the response: 7/10

Human evaluation of open domain dialog systems

The common practice is to use 1-turn evaluation

Context:

I think I want an ipad2. but I already have an ibook. Want vs. need.

Response:



Abuman	[USER] Sould I buy an ipad2?						
A numan rates how	i'm sorry, i'm sorry. i don't want to know.						
good the	[USER] vou're weird						
response:	you're welld.						
7/10	i'm sorry.						

Evaluating Dialog Systems: Our Proposal

- We need interactive evaluation
- We need to evaluate different aspects of the system
- **Quality**: overall, how was the quality of the chat?
- **Diversity**: how non-repetitive were the chat bot's responses?
- **Fluency**: how correct were the grammar and sentence structure of the chat bot's response?
- **Contingency**: how related to your messages were the chat bot's response?
- **Empathy**: how emotionally appropriate were the chat bot's response?

Our open source platform for Multi-turn evaluation



Dialog Models Generative Neural Network Models

HRED VHRED VHCR

References:

[1] Iulian V Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[2] Iulian Vlad Serban, Alessandro Sordoni, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron Courville, and Yoshua Bengio. A hierarchical latent variable encoder-decoder model for generating dialogues. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[3] Yookoon Park, Jaemin Cho, and Gunhee Kim. A hierarchical latent structure for variational conversation modeling. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1792–1801, 2018.

Dialog Models Generative Neural Network Models

	VHCR + EI

Regularization technique that makes the dialog model to be more aware of:

- The text sentiment
- The text topic







[1] A. Conneau, D. Kiela, H. Schwenk, L. Barrault, and A. Bordes. Super-vised learning of universal sentence representations from natural language inference data. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 670–680, 2017.

Data

Cornell Dataset

- The largest and most commonly used movie scripts dataset
- ~ 200K conversational exchanges between pairs of characters engaging in at least 3 turns

Reddit Dataset

- Conversations forum
- 109K conversations of at least 3 turns from conversational exchanges on the platform in 2018

Human Interactive Evaluation

		Cor	nell	Red	ldit
Model	Metric	Baseline	EI	Baseline	EI
	quality	2.182 ± 0.305	2.347 ± 0.313	2.527 ± 0.310	2.714 ± 0.299
HRED	fluency	3.909 ± 0.387	4.000 ± 0.381	4.436 ± 0.349	4.786 ± 0.316
	diversity	2.836 ± 0.374	2.735 ± 0.380	3.418 ± 0.386	3.554 ± 0.372
	contingency	2.200 ± 0.291	2.469 ± 0.336	2.382 ± 0.288	2.536 ± 0.322
	empathy	2.673 ± 0.352	2.490 ± 0.350	3.018 ± 0.329	3.107 ± 0.337
VHRED	quality	2.022 ± 0.309	2.333 ± 0.252	2.694 ± 0.392	2.864 ± 0.341
	fluency	3.109 ± 0.351	3.949 ± 0.396	4.250 ± 0.496	4.477 \pm 0.402
	diversity	3.565 ± 0.442	4.385 ± 0.371	$\textbf{5.00} \pm 0.468$	4.705 ± 0.353
	contingency	2.261 ± 0.287	2.487 ± 0.346	2.472 ± 0.362	2.773 ± 0.370
	empathy	2.739 ± 0.374	2.564 ± 0.367	3.000 ± 0.393	$\textbf{3.341} \pm 0.385$
	quality	2.132 ± 0.247	2.548 ± 0.380	2.615 ± 0.350	2.692 ± 0.298
	fluency	2.679 ± 0.306	3.976 ± 0.380	3.923 ± 0.433	4.308 ± 0.395
VHCR	diversity	3.755 ± 0.340	4.238 ± 0.421	4.436 ± 0.455	4.231 ± 0.382
	contingency	2.189 ± 0.270	2.571 ± 0.356	2.077 ± 0.298	2.692 ± 0.354
	empathy	2.340 ± 0.316	2.714 ± 0.368	2.974 ± 0.434	$\textbf{3.288} \pm 0.379$

Traditional Automatic Metrics

			(Cornell		Reddit					
Model	Version	PPL	KL	Avg	Ext	Grd	PPL	KL	Avg	Ext	Grd
HPED	baseline	52.311	11 13	.471	.329	.331	41.730	-	.649	.394	.474
IKED	EI	47.636	-	.560	.383	.400	41.245	-	.651	.398	.482
VHPED	baseline	49.414	.264	.539	.352	.395	36.240	.188	.635	.383	.464
VIIKED	EI	50.526	.517	.545	.355	.394	35.510	.167	.636	.392	.465
VHCP	baseline	61.000	.562	.532	.345	.382	36.736	.267	.619	.371	.448
VIICK	EI	49.243	.475	.588	.369	.444	37.198	.231	.639	.394	.469

Examples of dialogs

[User]: I would love to make a trip! Where would you go?
<pre>[Chatbot]: not sure about it.</pre>
[User]: I see. I would go to Mexico
Chatbot]: why?
[User]: because it's beautiful and warm
[Chatbot]: to be honest i haven't been to a lot of countries, but i don't really have a lot of fun
[User]: I think you could virtually travel to any place in the world!
<pre>[Chatbot]: '+ it's a great way to make new friends.</pre>
[User]: That's true!
<pre></pre>
[User]: sometimes you have to choose between two people, it's true
[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome.</unknown>
[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a ≺unknown≻, you're pretty awesome. [User]: thanks, you are awesome too!
[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: 🍐 🐵 🍪 do you like music?</unknown>
[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown≻, awesome.<br="" pretty="" you're="">[User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: de ee ee ee do you like music? [User]: I love music</unknown≻,>
<pre>[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: de ee ee</unknown></pre>
<pre>[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: ▲</unknown></pre>
<pre>[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: ▲ ④ ④ ⑤ do you like music? [User]: I love music [ELSA-emo VCHR reddit]: I · i can't imagine how life would be without it. [User]: haha I agree [ELSA-emo VCHR reddit]: ▲ ③ i m not sure if you re joking or not, i m not sure if you re joking or not.</unknown></pre>
<pre>[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: ▲ ④ ④ ④ ❺ do you like music? [User]: I love music [ELSA-emo VCHR reddit]: ♬ ☆ i can't imagine how life would be without it. [User]: haha I agree [ELSA-emo VCHR reddit]: ▲ ④ i m not sure if you re joking or not, i m not sure if you re joking or not. [User]: I'm not joking. Life without music would be so horrible</unknown></pre>
<pre>[User]: sometimes you have to choose between two people, it's true [ELSA-emo VCHR reddit]: like you said, you're a <unknown>, you're pretty awesome. [User]: thanks, you are awesome too! [ELSA-emo VCHR reddit]: d e e e o o do you like music? [User]: I love music [ELSA-emo VCHR reddit]: f → i can't imagine how life would be without it. [User]: haha I agree [ELSA-emo VCHR reddit]: d o i m not sure if you re joking or not, i m not sure if you re joking or not. [User]: I'm not joking. Life without music would be so horrible [ELSA-emo VCHR reddit]: f → i love music too!</unknown></pre>

1. Sentiment Metrics

Sentiment Score; Sentiment Coherence; Sentiment Transition; Sentiment Min-Max; Laughter

2. Engagement Metrics

Question Score; #Words

3. Semantic Metrics

Semantic Similarity; Average Word Coherence; Extrema Word Coherence; Greedy Word Coherence

Hybrid Metric

M1, M2, ..., M11

$$M_H = \sum \lambda_i * M_i + M_0$$

1. Sentiment Metrics

Sentiment Score; Sentiment Coherence; Sentiment Transition; Sentiment Min-Max; Laughter

2. Engagement Metrics

Question Score; #Words

3. Semantic Metrics

Semantic Similarity; Average Word Coherence; Extrema Word Coherence; Greedy Word Coherence Goal: to approximate the human ratings

Hybrid Metric

M1, M2, ..., M11

$$M_H = \sum \lambda_i * M_i + M_0$$

Multi-turn HUMAN-BOT conversation EI-HRED trained on Reddit

[User] talk to me [EI HRED]: it's my job. [User]: i love you [EI HRED]: thank you for the kind words. [User]: do you love me? [EI HRED]: it's a bit personal, I sent you a dm. [User]: i'm blushing

Hybrid Metric

$$M_H = \sum \lambda_i * M_i + M_0$$

Human evaluation

High correlation

Multi-turn HUMAN-BOT conversation EI-HRED trained on Reddit

[User] talk to me [EI HRED]: it's my job. [User]: i love you [EI HRED]: thank you for the kind words. [User]: do you love me? [EI HRED]: it's a bit personal, I sent you a dm. [User]: i'm blushing

Hybrid Metric

$$M_H = \sum \lambda_i * M_i + M_0$$

Self-Play Scenario

Multi-turn HUMAN-BOT conversation	Multi-turn BOT-BOT conversation					
EI-HRED trained on Reddit	EI-HRED trained on Reddit					
[User] talk to me	[EI HRED] i can't wait to see what					
[EI HRED]: it's my job.	i can do.					
[User]: i love you	[EI HRED] that's a good idea.					
[EI HRED]: thank you for the kind words.	[EI HRED] how do you deal with people?					
[User]: do you love me?	[EI HRED] i've been on reddit					
[EI HRED]: it's a bit personal, I sent you a dm.	for a long time.					
[User]: i'm blushing	[EI HRED] i am in the same boat.					

Hybrid Metric

$$M_H = \sum \lambda_i * M_i + M_0$$

Self-Play Scenario

Multi-turn BOT-BOT conversation EI-HRED trained on Reddit

[EI HRED] i can't wait to see what i can do. [EI HRED] that's a good idea.

[EI HRED] how do you deal with people?

[EI HRED] i've been on reddit

for a long time.

[EI HRED] i am in the same boat.

Correlation between Metrics and Human Interactive Evaluation

Traditional Automatic Metrics

New Automatic Metrics (on self play scenario)

		L	Au	tomatic I	M.		Sentiment M.						Semantic M.				Engagement M.		
arion	Quality	-0.145	-0.145	0.141	0.100	0.132	0.214	-0.089	-0.038	0.044	0.069	-0.010	0.035	-0.007	0.042	0.013	0.048	0.725	- 0.8
EVAIN	Diversity	-0.167	-0.157	0.166	0.078	0.171	0.145	-0.050	-0.017	-0.068	0.053	-0.005	0.024	-0.052	0.056	0.099	0.000	0.640	- 0.4
Idcuve	Fluency	- 0.204	-0.212	0.188	0.167	0.180	0.173	0.023	-0.050	0.017	0.022	0.065	0.127	0.063	0.076	-0.007	0.070	0.455	- 0.0
	Contingency	0.046	-0.047	0.059	0.079	0.060	0.177	-0.083	-0.051	0.034	0.034	-0.044	0.033	0.030	0.029	0.005	0.039	0.263	0.4
	Empathy	0.178	-0.177	0.152	0.116	0.144	0.264	-0.093	-0.060	-0.055	0.103	-0.079	0.015	0.040	0.067	0.032	-0.070	0.834	0.8
	Bitsp	erword per	plexici p	yetage 6	thema.	steedy' sentir	sent Trans	Sent Min	hat U Laud	Sent. Coh	enantic Coh	er.UB	raword Cor	er.UB	Question	core 8 * W	id meric m	AL BIB	



PAPER:

Asma Ghandeharioun*, Judy Hanwen Shen*, Natasha Jaques*, Craig Ferguson, Noah Jones, Agata Lapedriza, Rosalind Picard, "Approximating Interactive Human Evaluation with Self-Play for Open-Domain Dialog", *in Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2019.

CODE: https://github.com/natashamjaques/neural_chat

Applications of Affective Computing

- Emotional Wellbeing
- User Experience
- Suicide prevention
- Early detection of depression
- Education

Emotional Wellbeing







Face Analysis



Detected Expressions: attention: 38.0

chinRaise: 8.0 dimpler: 2.0 lipSuck: 43.0 noseWrinkle: 2.0 smile: 47.0 smirk: 37.0 **Body Analysis**



Detected Body Parts:

15. Right Eye	18. Left Ear	5. L. Shoulder				
0. Nose	1. Neck	3. Right Elbow				
16. Left Eye	2. R. Shoulder	4. Right Wrist				

Place Analysis



Place Category: Indoor, Classroom (0,188) Place Attributes: No-horizon, enclosed area, man-made, cloth, working, indoor lighting, paper, wood, studying

Stress-free Driving

Thank you to our partners



Javier

Hernandez

Diego

Muñoz

Emotional Navigation SIG, enavigation.media.mit.edu Sponsors: Hyundai, NTT Data, Daimler Multimodal Sensing
-Heart Rate
-Respiration
-Pressure sensors on
the steering wheel
-Face Analysis
-Context Analysis
Car interventions for
a better driving

experience.

Craig

Ferguson

Vincent

Chen

Final Conclusions

Emotion AI technologies are useful and have a lot of applications.

We are nowhere close to have machines that understand emotions as humans do.

We are slowly finding solutions to specific problems that require Emotional Intelligence.
References and Resources

Emotions in Context http://sunai.uoc.edu/emotic/

ا 🐴 😖 💽

[1] R. Kosti, J.M Alvarez, A. Recasens, A.Lapedriza. "Emotion Recognition in Context". Computer Vision and Pattern Recognition (CVPR), 2017

[2] R. Kosti, J.M Alvarez, A. Recasens, A.Lapedriza. "Context based Emotion Recognition using EMOTIC dataset". IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2019.

places



http://places2.csail.mit.edu

[3] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba. "Places: A 10 Million Image Database for Scene Recognition". IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), July 2017.

[4] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. "Object Detectors Emerge in Deep Scene CNNs." International Conference on Learning Representations (ICLR), 2015.

CAM (Class Activation Maps)



http://cnnlocalization.csail.mit.edu

[5] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. "Learning Deep Features for Discriminative Localization". Computer Vision and Pattern Recognition (CVPR), 2016.

Sequence Bias

Universitat

Oberta



[6] Judy Shen, Agata Lapedriza, Rosalind Picard, "Unintentional affective priming during labeling may bias labels", 8th International Conference on Affective Computing and Intelligent Interaction, Cambridge, UK, 2019.

Dialog & Neural Chat



https://www.media.mit.edu/ projects/elsa/overview/

[7] Asma Ghandeharioun*, Judy Hanwen Shen*, Natasha Jaques*, Craig Ferguson, Noah Jones, Agata Lapedriza, Rosalind Picard, "Approximating Interactive Human Evaluation with Self-Play for Open-Domain Dialog", NeurIPS, Vancouver, Canada, 2019