Emotion Detection in Text

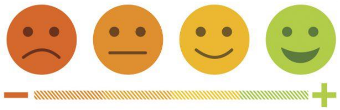
by Arjun Nair (10/12/21)

What is emotion?

- There is no exact scientific consensus on an exact definition of emotion
- Generally characterized as something that is:
  - A direct response to some event (either internal or external) (Schachter et al., 2011)
    - Internal: Some thought that the person had
    - External: Something that happened in the outside world
  - Takes the form of neurotransmitters and hormones (Wang et al., 2020)
  - Short-term (as opposed to mood) (Ekman, 1999)
    - Some disagreement on this (for example, is grief an emotion?)
  - Associated with either pleasure or displeasure (Wilson et al., 2004)
    - In the next slide, we will discuss how emotion is different from sentiment
  - Characterized by the way that humans tend to act when experiencing them (Fox, 2008)
- The term **affect** is used to refer more broadly to emotion, mood, and feeling
  - As such, emotion detection is often referred to as **affective computing**

Emotion v.s. Sentiment

- Sentiment is strictly about **positive vs. negative**
- Emotion is more specific than sentiment
  - For example, **sadness** and **anger** are both negative emotions; however, they are fundamentally different in terms of the actions that humans take when experiencing them
  - **Sadness**: When experiencing this emotion, people tend to respond passively and direct their negativity inwards; oftentimes, people who are sad report that they experience feelings of helplessness
  - **Anger**: Humans experiencing this emotion tend to direct their negativity outwards and lash out in some way - either physically or verbally



Sentiment

Emotion

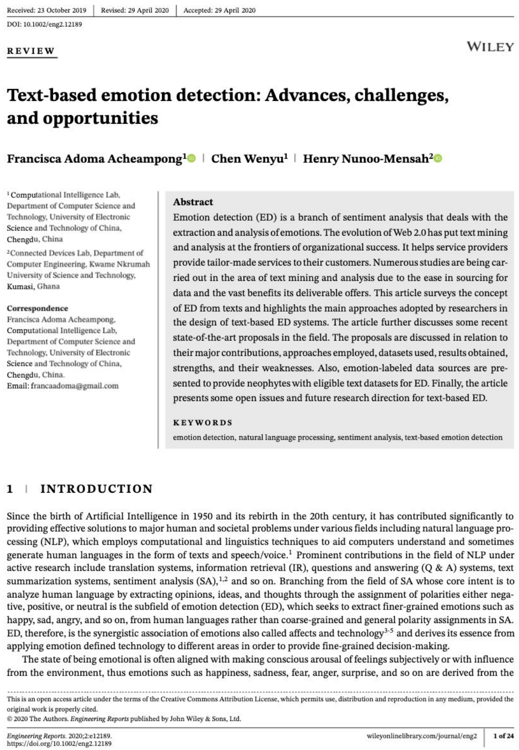
Emotion Detection (Affective Computing)

- Detecting a person's emotion (or, more broadly, affective state) at a given time, either from:
  - Text (NLP)
    - Focus of this presentation
  - Audio (Digital Signal Processing + NLP)
  - Body Language (Computer Vision)
    - Facial Expressions
    - Posture
    - Hand Gestures
- Utilizes machine learning techniques
  - Typically modeled as a regression or classification problem
  - Output depends on which model of emotion you are using

[Required Paper] Text-based emotion detection: Advances, challenges, and opportunities

- Survey paper that covers the state of the field as of May 2020
- Does a really good job of summarizing:
  - Models of Emotion
  - Important Datasets

Authors: Francisca Adoma Acheampong, Chen Wenyu, Henry Nunoo-Mensah
Venue: Engineering Reports
Year: 2020



Received: 23 October 2019 | Revised: 29 April 2020 | Accepted: 29 April 2020
DOI: 10.1002/eng2.12189

REVIEW

WILEY

Text-based emotion detection: Advances, challenges, and opportunities

Francisca Adoma Acheampong[1] | Chen Wenyu[1] | Henry Nunoo-Mensah[2]

[1] Computational Intelligence Lab, Department of Computer Science and Technology, University of Electronic Science and Technology of China, Chengdu, China
[2] Connected Devices Lab, Department of Computer Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

Correspondence
Francisca Adoma Acheampong, Computational Intelligence Lab, Department of Computer Science and Technology, University of Electronic Science and Technology of China, Chengdu, China.
Email: francaadoma@gmail.com

Abstract
Emotion detection (ED) is a branch of sentiment analysis that deals with the extraction and analysis of emotions. The evolution of Web 2.0 has put text mining and analysis at the frontiers of organizational success. It helps service providers provide tailor-made services to their customers. Numerous studies are being carried out in the area of text mining and analysis due to the ease in sourcing for data and the vast benefits its deliverable offers. This article surveys the concept of ED from texts and highlights the main approaches adopted by researchers in the design of text-based ED systems. The article further discusses some recent state-of-the-art proposals in the field. The proposals are discussed in relation to their major contributions, approaches employed, datasets used, results obtained, strengths, and their weaknesses. Also, emotion-labeled data sources are presented to provide neophytes with eligible text datasets for ED. Finally, the article presents some open issues and future research direction for text-based ED.

KEYWORDS
emotion detection, natural language processing, sentiment analysis, text-based emotion detection

1 | INTRODUCTION

Since the birth of Artificial Intelligence in 1950 and its rebirth in the 20th century, it has contributed significantly to providing effective solutions to major human and societal problems under various fields including natural language processing (NLP), which employs computational and linguistics techniques to aid computers understand and sometimes generate human languages in the form of texts and speech/voice.[1] Prominent contributions in the field of NLP under active research include translation systems, information retrieval (IR), questions and answering (Q & A) systems, text summarization systems, sentiment analysis (SA),[1,2] and so on. Branching from the field of SA whose core intent is to analyze human language by extracting opinions, ideas, and thoughts through the assignment of polarities either negative, positive, or neutral is the subfield of emotion detection (ED), which seeks to extract finer-grained emotions such as happy, sad, angry, and so on, from human languages rather than coarse-grained and general polarity assignments in SA. ED, therefore, is the synergistic association of emotions also called affects and technology[3-5] and derives its essence from applying emotion defined technology to different areas in order to provide fine-grained decision-making.

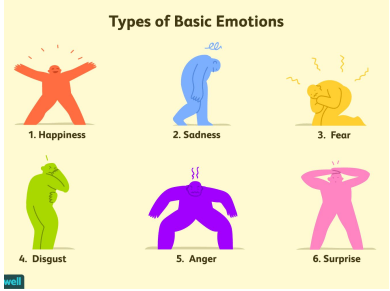
The state of being emotional is often aligned with making conscious arousal of feelings subjectively or with influence from the environment, thus emotions such as happiness, sadness, fear, anger, surprise, and so on are derived from the

Models of Emotion: Introduction

- Invented by psychologists and then used by computer scientists to model their regression or classification problem
- Two types of models of emotion
  - **Discrete** (or categorical) models place emotions into distinct classes or categories
    - Modeled as a classification problem
  - **Dimensional** models of emotion place emotions into a coordinate space with different axes
    - Makes the assumption that emotions can be related to each other
    - Distance between two emotions on this space will show how similar they are
    - Most modern dimensional models are multidimensional (two or more dimensions)
    - Modeled as a regression problem with one output per dimension
- The paper claims that discrete models are the most commonly used in emotion detection; however, the source it cites (Canales et al., 2014) in turn cites another paper (Calvo et al., 2013) that does not provide evidence for this claim
  - Based off of personal experience, I think dimensional models of emotion are more common

Ekman’s Basic Emotions (Discrete)

- Psychologist Paul Ekman proposed six basic emotions derived from universal facial expressions (Ekman, 1970)
  - Happiness
  - Sadness
  - Anger
  - Disgust
  - Fear
  - Surprise
- Synergy of these emotions may produce more complex emotions such as guilt, shame, pride, greed, etc. (Ekman and Freisen, 1975)
  - In a 1999 book, Ekman said that he is no longer sure whether this is true (Ekman, 1999)
- Ekman’s lab has also shown strong evidence of a seventh basic emotion: contempt (Ekman, 1988)
  - This has not been universally adopted by the scientific community and is uncommon in the field of emotion detection



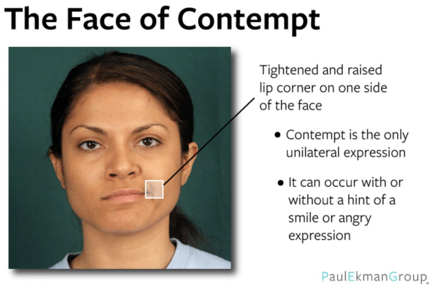


Table 3.1 Characteristics which distinguish basic emotions from one another and from other affective phenomena

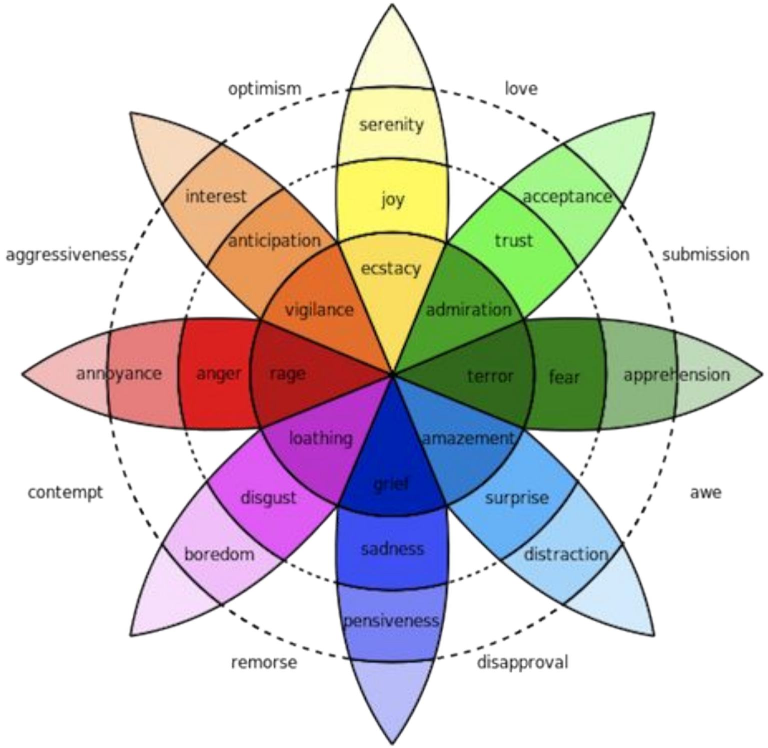
1. Distinctive universal signals
2. Distinctive physiology
3. Automatic appraisal, tuned to:
4. Distinctive universals in antecedent events
5. Distinctive appearance developmentally
6. Presence in other primates
7. Quick onset
8. Brief duration
9. Unbidden occurrence
10. Distinctive thoughts, memories images
11. Distinctive subjective experience

Ekman’s Basic Emotions (Dimensional)

- Even though Ekman proposed a discrete model (each emotion either exists or it doesn’t), emotion detection researchers have adapted it into a 6-dimensional model where each basic emotion is a separate axis
  - Oftentimes, sentiment (usually referred to as **valence**) is added as a seventh dimension
- The ML problem is often framed as the following
  - **Input**: Text, either as a bag of words, n-grams, or a sequence of tokens
  - **Output**: 6-tuple with each member being in the range [0,1] (ex. [0.55, 0.32, 0.44…]
- Used by several emotion detection systems
  - ex. IBM Watson NLU

Plutchik's Wheel of Emotions (Dimensional)

- Robert Plutchik posited that there were eight emotions in opposition to each other (Plutchik, 1980)
  - Aggressiveness v.s. Awe
  - Submission v.s. Contempt
  - Optimism v.s. Disapproval
  - Remorse v.s. Love
- This has been adapted into a two-dimensional "wheel" model
  - The circular axis is where the emotion pairs lie
  - The depth is the **intensity** of the emotion



Ortony, Clore, and Collins (OCC) Model (Discrete)

- The OCC model postulates that there are 22 different emotions, including Ekman's six basic emotions (Ortony et al., 1988)
  - Additional emotions include relief, envy, reproach, self-reproach, appreciation, shame, pity, disappointment, admiration, hope, fears-confirmed, grief, gratification, gloating, like, and dislike
- This has been used before in emotion detection (Jiang et al., 2016; Shaikh et al., 2009) but is not nearly as popular as Ekman's model
- There are a number of ambiguities in the model that make it difficult to use in a computational context (Steunebrink et al., 2009)

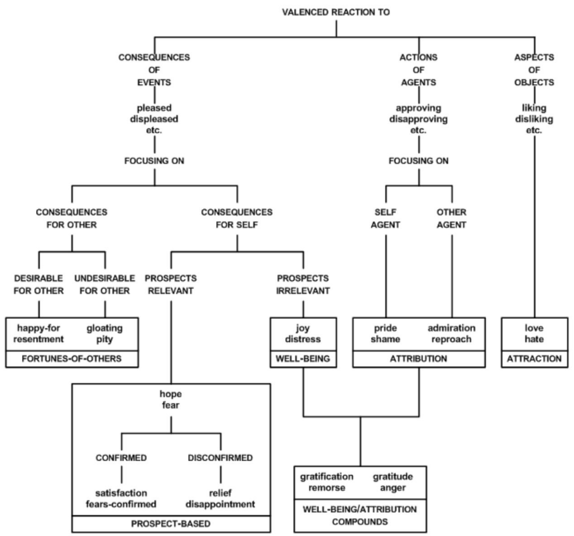
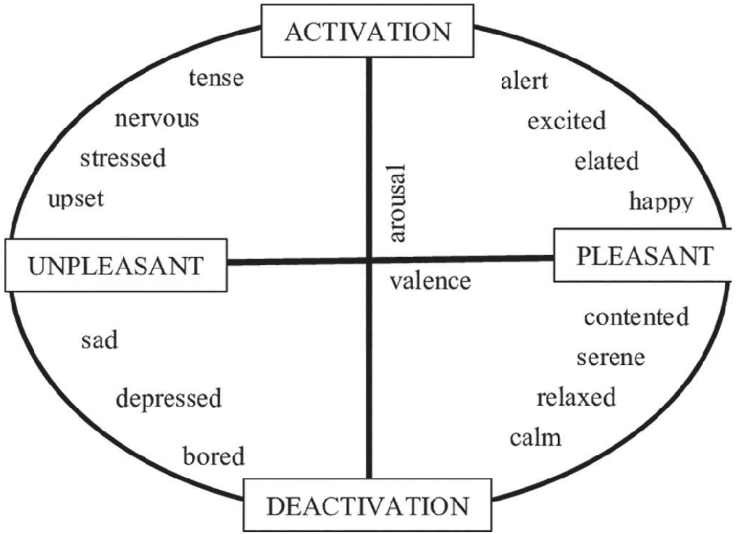


Fig. 1. The original structure of emotions of the OCC model, copied from page 19 [1].

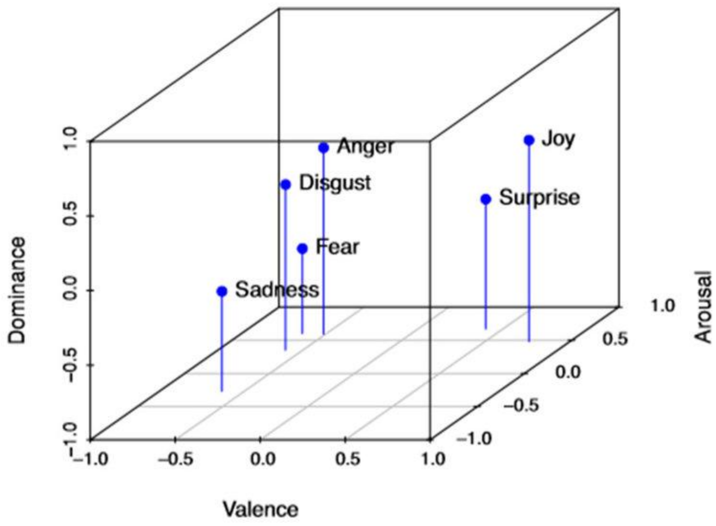
Circumplex of Affect (Dimensional)

- 2-dimensional model with **arousal** and **valence** as the axes (Russell, 1980)
  - **Arousal** represents how "activated" a person is when experiencing the emotion
  - **Valence** (i.e. sentiment) is how positive or negative something is



Three-Factor Theory of Emotions (VAD)

- Russel and Mehrabian (1977) proposed an additional axis called **dominance** that represents an individual's sense of control over their surroundings
- Emotion (and emotion detection) research has largely ignored this additional dimension; however, fMRI scans have shown the same types of neural correlates with dominance that are seen with valence and arousal (Jerram et al., 2014)



Datasets

- Since emotion detection systems typically utilize established regression or classification models, the most difficult aspect of emotion detection is the collection of reliable, labeled data
  - Semi-supervised learning can reduce (but not eliminate) the need for labels
- Typically this is done via social media (Twitter, Facebook, etc.)
  - Use some correlated feature, like an emoji or a hashtag, and assume emotion based off of that
  - Or annotate (manual, crowdsourcing, etc.)
- We will talk more about datasets later in the presentation

Discussion

- What do you think about the models of emotion discussed during this far? Are there any that you prefer in the context of ML?
    - Ekman’s six basic emotions, VAD (valence-arousal-dominance), Plutchik’s wheel of emotions, OCC model, circumplex of affect (valence-arousal)

[Required Paper] Multimodal and Multi-view Models for Emotion Recognition

- **Mode**: Medium of communication
  - Text, image, audio, etc.
- **Multimodal**: Combine multiple modes
  - Picture book, video with sound, etc.
- **Multiview Model**: A type of ML model that integrates two or more “views” (or representations) of the data
- **Contrastive Learning**: Uses negative samples to show the model what *not* to output

Multimodal and Multi-view Models for Emotion Recognition

Gustavo Aguilar[‡], Viktor Rozgić[†], Weiran Wang[†] and Chao Wang[†]
University of Houston[‡]
Amazon.com[†]
gaguilaralas@uh.edu, {rozgicv, weiranw, wngcha}@amazon.com

arXiv:1906.10198v1 [cs.CL] 24 Jun 2019

Abstract

Studies on emotion recognition (ER) show that combining lexical and acoustic information results in more robust and accurate models. The majority of the studies focus on settings where both modalities are available in training and evaluation. However, in practice, this is not always the case; getting ASR output may represent a bottleneck in a deployment pipeline due to computational complexity or privacy-related constraints. To address this challenge, we study the problem of efficiently combining acoustic and lexical modalities during training while still providing a deployable acoustic model that does not require lexical inputs. We first experiment with multimodal models and two attention mechanisms to assess the extent of the benefits that lexical information can provide. Then, we frame the task as a multi-view learning problem to induce semantic information from a multimodal model into our acoustic-only network using a contrastive loss function. Our multimodal model outperforms the previous state of the art on the USC-IEMOCAP dataset reported on lexical and acoustic information. Additionally, our multi-view-trained acoustic network significantly surpasses models that have been exclusively trained with acoustic features.

1 Introduction

The task of emotion recognition (ER) requires understanding the way humans interact to express their emotional state during conversations. Among others, emotions are encoded in both lexical and acoustic information where each modality contributes to the overall emotional state of a given speaker. However, in some situations, one modality can be more insightful to derive emotions than the other. For instance, the phrase *“yeah... of course”* does not have enough lexical information to derive the right emotion, and it may all depend on the acoustic patterns. On the other hand, the phrase *“I really miss my dog!”* does not need acoustic information to detect that the most likely emotion is sadness. Thus, recognizing emotions is not a trivial task because an emotional state can be easily shaped by many factors: context, word content, spectral and prosodic information, among others (Barbulescu et al., 2017).

In this paper, we study the emotion recognition problem from the speech and language perspectives. We formally look into acoustic and lexical modalities with the aim of improving models that only use acoustic information. In the first part of this work, our goal is to assess the extent to which lexical information benefits acoustic models. We propose a multimodal method that is inspired by the way humans process emotions in a conversation. That is, lexical and acoustic information is simultaneously perceived at every word step. Hence, we introduce the concept of acoustic words: word-level representations derived from acoustic features in a speech fragment. The acoustic word representations enable a natural combination of the modalities where lexical and acoustic features are aligned at the word level. Additionally, we leverage these representations with two attention mechanisms: modality-based and context-based attentions. The former mechanism prioritizes one of the modalities at each word step, whereas the latter mechanism focuses on the most important word representations across the entire utterance. Our multimodal approach outperforms the current state of the art on the USC-IEMOCAP dataset reported on lexical and acoustic modalities.

In the second part of this work, our goal is to induce semantic information from the proposed multimodal model into an acoustic model. We study a more challenging scenario where we establish that lexical information is available during

Authors: Gustavo Aguilar, Viktor Rozgic, Weiran Wang, Chao Wang
Venue: ACL
Year: 2019

Introduction

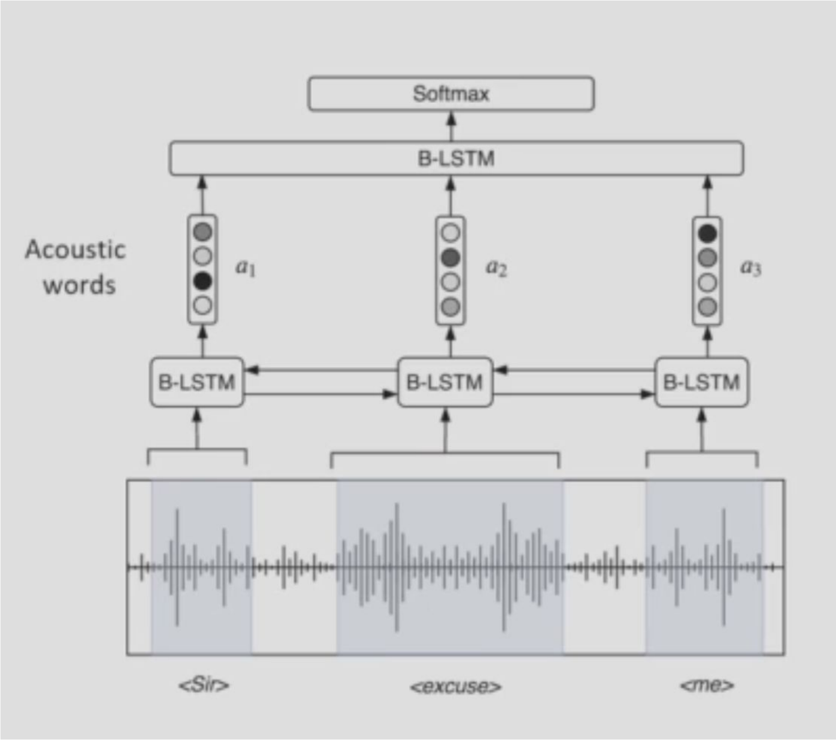
- Previous studies have shown that integrating acoustic and lexical information together into a multimodal system can improve the performance of emotion detection systems (Jin et al., 2015)
- It is often challenging to get lexical data to accompany acoustic data
  - Requires annotators to convert the speech to text
  - Or an automated speech recognition (ASR) system, which may be buggy or slow
- The aim of this paper is to develop a unified multimodal, multi-view model that can take either audio only or audio/lexical input depending on what is available

IEMOCAP Dataset

- The Interactive Emotional Dyadic Motion Capture dataset is a multispeaker, multimodal database created by the USC SAIL Lab ([Busso et al., 2008](#))
- Contains twelve hours of audiovisual data, including:
    - Video
    - Speech
    - Motion capture of face
    - Text transcriptions
- Annotated by multiple annotators into:
    - Categorical labels: anger, happiness, sadness, and neutrality
    - Dimensional labels: VAD (valence-arousal-dominance)
- The model for this paper takes the speech data as input and outputs a prediction for the categorical label
    - However, during training, it also uses text transcription (lexical) data as input

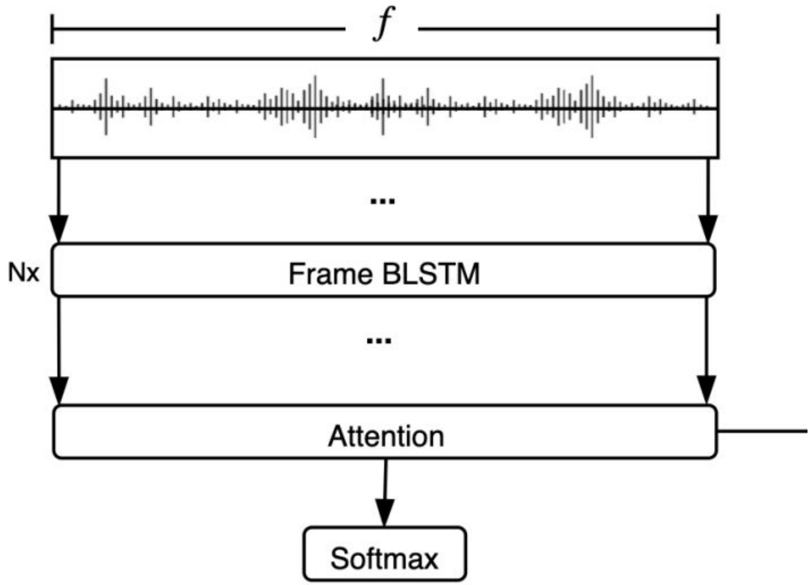
Acoustic Words

- Previous studies on multimodal NLP have taken acoustic representations and concatenated them with lexical representations before feeding them into a model ([He et al., 2016](#))
- The authors argue that a more natural way to represent the data is to instead align words from the acoustic features ("acoustic words") with the words in the lexical features
    - Use attention to determine whether the lexical or acoustic word contributes more to the emotional content of the utterance
- Allows the model to fuse acoustic and lexical modalities together during training



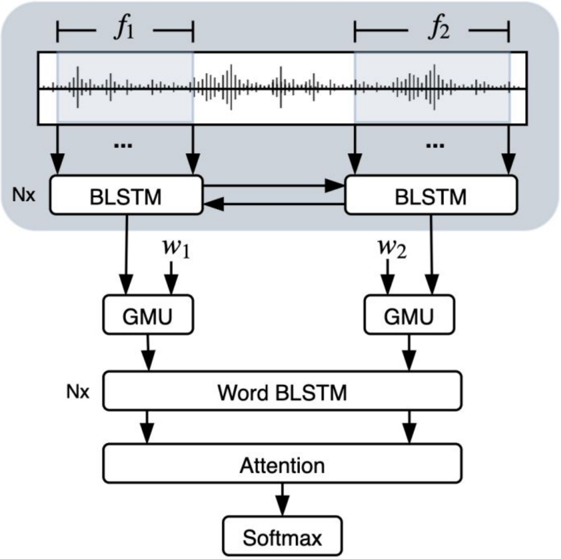
View 1: Acoustic Model

- Uses OpenSMILE (Eyben et. al, 2013) to convert audio into frame-level features
  - Examples: energy, MFCC, etc.
- A **BLSTM** (Bidirectional Long Short-Term Memory Network) takes those features and produces an output that gets fed into an attention layer

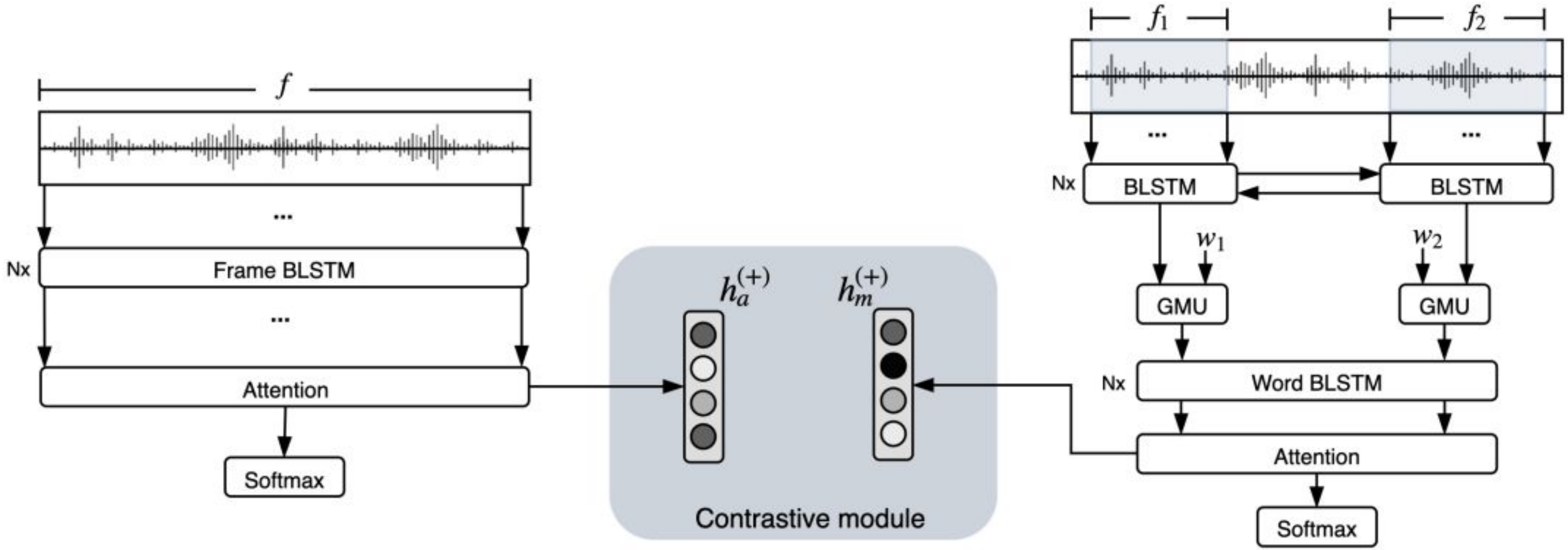


View 2: Multimodal Model (Acoustic + Lexical)

- Uses BLSTM to take frame-level features and output acoustic representations
- Feeds acoustic word representation along with its corresponding lexical word into a GMU
    - **GMU** (Gated Multimodal Unit): Decides which input contributes more to the emotional content
- Output of the GMUs are fed into another BLSTM that gets fed into an attention layer



Multiview Model: View 1 + View 2



Results

Type	Experiment	Modality	Dev	Test	Comment
Baseline	B-ACO-1	Acoustics	0.5858	-	Silence frames
	B-ACO-2		0.5729	-	Silence frames removed
	B-LEX	Lexical	0.6706	-	-
	B-MM-1	Multimodal	0.7195	-	Silence frames
	B-MM-2		0.7265	-	Silence frames removed
Hierarchical	H-ACO-1	Acoustics	0.5697	-	Acoustic words
	H-MM-1	Multimodal	0.7316	-	Aligned words
	H-MM-2		0.7341	-	+ GMU
	H-MM-3		0.7354	-	+ Attention
	H-MM-4		**0.7383**	**0.7169**	+ GMU + Attention
SOTA	-	Multimodal	-	**0.7079**	Poria et al. (2018)

Table 2: The results of the multimodal experiments. The name of the experiments starts either with B or H referring to baseline or hierarchical models. ACO, LEX, and MM mean acoustic, lexical and multimodal. Our results provide a new state-of-the-art UA when we use the hierarchical model with GMU and attention. Once the models are optimized on the validation set, we evaluate the best ones on the test set.

Discussion

- What do you think of the alignment of acoustic word representations with the lexical words that they correspond to?
- Why do you think the performance of the acoustic model dropped as part of this multiview framework?

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

- Rather than detecting emotions directly, DeepMoji detects emojis, which serve as a rough proxy for affective state
- Data is pulled from Twitter, where the raw text of the tweet is used as input and the corresponding emoji is used as a label
    - If multiple different emojis show up in the same tweet, a copy is made for each of them

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

Bjarke Felbo[1], Alan Mislove[2], Anders Søgaard[3], Iyad Rahwan[1], Sune Lehmann[4]

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

arXiv:1708.00524v2 [stat.ML] 7 Oct 2017

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-the-art 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 analysis 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 extending 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 sarcasm. 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 love mom's cooking	49.1%	8.8%	3.1%	3.0%	2.9%
I love how you never reply back..	14.0%	8.3%	6.3%	5.4%	5.1%
I love cruising with my homies	34.0%	6.6%	5.7%	4.1%	3.8%
I love messing with yo mind!!	17.2%	11.8%	8.0%	6.4%	5.3%
I love you and now you're just gone..	39.1%	11.0%	7.3%	5.3%	4.5%
This is shit	7.0%	6.4%	6.0%	6.0%	5.8%
This is the shit	10.9%	9.7%	6.5%	5.7%	4.8%

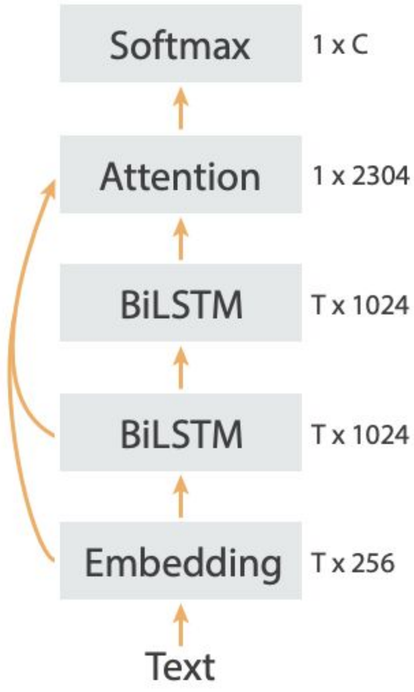
Emojis are not always a direct labeling of emotional content. For instance, a positive emoji may serve to disambiguate an ambiguous sentence or to complement an otherwise relatively negative text. Kunneman et al. (2014) discuss a similar duality in the use of emotional hashtags such as *#nice* and *#lame*. 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

Authors: Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, Sune Lehmann
Venue: EMNLP
Year: 2017

Pretraining and Model Architecture

- In order to ensure that the models learn emotional content for all emojis rather than just the most used ones, the authors:
  - Create a balanced pretraining dataset from the original data
  - Fine-tune using the rest of the data
    - Upsampled to create a balanced fine-tuning dataset
- Model architecture utilizes:
  - A 256-node embedding layer
  - Two 1024-node BiLSTM layers
  - 2304-node attention layer
    - With residual (skip connections) from previous layers
  - Softmax layer that outputs probabilities for each emoji



Chain-Thaw Fine-Tuning

- Fine-tunes layers one-by-one and then finally fine-tunes all layers together
- Expands vocabulary to new domains (such as science, film, etc.) without overfitting the model
- Has not been used in another paper ever since then
  - Except arguably [Zunair et al., 2018], which DeepMoji author Bjarke Felbo [claims uses the same approach] (but others disagree)

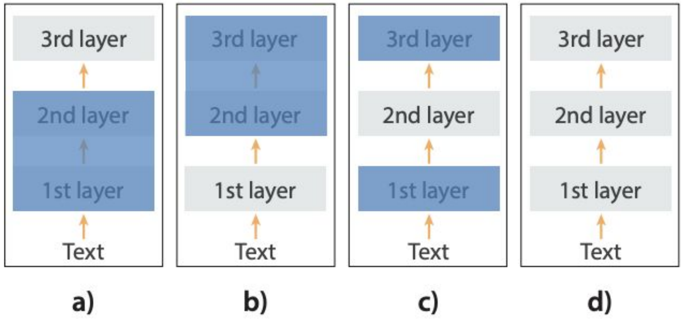


Figure 2: Illustration of the chain-thaw transfer learning approach, where each layer is fine-tuned separately. Layers covered with a blue rectangle are frozen. Step a) tunes any new layers, b) then tunes the 1st layer and c) the next layer until all layers have been fine-tuned individually. Lastly, in step d) all layers are fine-tuned together.

Results

Table 4: Description of benchmark datasets. Datasets without pre-existing training/test splits are split by us (with splits publicly available). Data used for hyperparameter tuning is taken from the training set.

Identifier	Study	Task	Domain	Classes	N_{train}	N_{test}
SE0714	(Strapparava and Mihalcea, 2007)	Emotion	Headlines	3	250	1000
Olympic	(Sintsova et al., 2013)	Emotion	Tweets	4	250	709
PsychExp	(Wallbott and Scherer, 1986)	Emotion	Experiences	7	1000	6480
SS-Twitter	(Thelwall et al., 2012)	Sentiment	Tweets	2	1000	1113
SS-Youtube	(Thelwall et al., 2012)	Sentiment	Video Comments	2	1000	1142
SE1604	(Nakov et al., 2016)	Sentiment	Tweets	3	7155	31986
SCv1	(Walker et al., 2012)	Sarcasm	Debate Forums	2	1000	995
SCv2-GEN	(Oraby et al., 2016)	Sarcasm	Debate Forums	2	1000	2260

Table 5: Comparison across benchmark datasets. Reported values are averages across five runs. Variations refer to transfer learning approaches in §3.3 with 'new' being a model trained without pretraining.

Dataset	Measure	State of the art	DeepMoji (new)	DeepMoji (full)	DeepMoji (last)	DeepMoji (chain-thaw)
SE0714	F1	.34 [Buechel]	.21	.31	.36	**.37**
Olympic	F1	.50 [Buechel]	.43	.50	**.61**	**.61**
PsychExp	F1	.45 [Buechel]	.32	.42	.56	**.57**
SS-Twitter	Acc	.82 [Deriu]	.62	.85	.87	**.88**
SS-Youtube	Acc	.86 [Deriu]	.75	.88	.92	**.93**
SE1604	Acc	.51 [Deriu][3]	.51	.54	**.58**	**.58**
SCv1	F1	.63 [Joshi]	.67	.65	.68	**.69**
SCv2-GEN	F1	.72 [Joshi]	.71	.71	.74	**.75**

Table 6: Benchmarks using a smaller emoji set (Pos/Neg emojis) or a classic architecture (standard LSTM). Results for DeepMoji from Table 5 are added for convenience. Evaluation metrics are as in Table 5. Reported values are the averages across five runs.

Dataset	Pos/Neg emojis	Standard LSTM	DeepMoji
SE0714	.32	.35	.36
Olympic	.55	.57	.61
PsychExp	.40	.49	.56
SS-Twitter	.86	.86	.87
SS-Youtube	.90	.91	.92
SE1604	.56	.57	.58
SCv1	.66	.66	.68
SCv2-GEN	.72	.73	.74

Dendrogram of Emojis

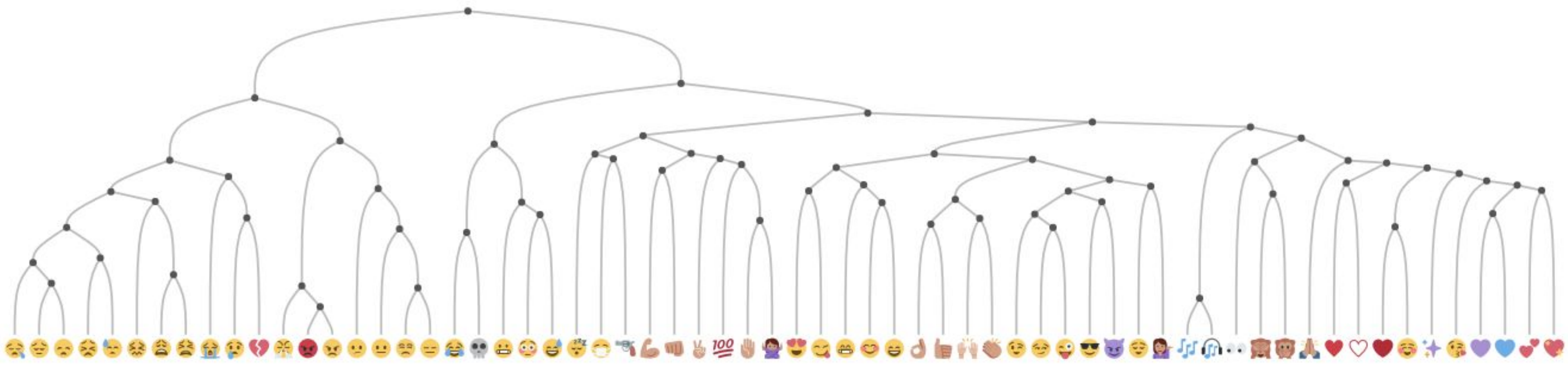


Figure 3: Hierarchical clustering of the DeepMoji model's predictions across categories on the test set. The dendrogram shows how the model learns to group emojis into overall categories and subcategories based on emotional content. The y-axis is the distance on the correlation matrix of the model's predictions measured using average linkage. More details are available in the supplementary material.

Discussion

- Are emojis a good proxy for affect?
- Any thoughts on the chain-thaw approach?
- DeepMoji was published on the same year as the first transformers paper. Do you think that using transformers instead of BiLSTMs could have yielded better performance on this task?

[Additional Paper] Modeling Empathy and Distress in Reaction to News Stories

- Aims to design a corpus for training empathy and distress detection systems
- Creates several (low-performing) regression models for this corpus, leaving the door open for other researchers to design better models

Modeling Empathy and Distress in Reaction to News Stories

Sven Buechel [*†3] Anneke Buffone [*1] Barry Slaff [1] Lyle Ungar [1,2] João Sedoc [1,2]

[1] Positive Psychology Center, University of Pennsylvania
[2] Computer & Information Science, University of Pennsylvania
[3] JULIE Lab, Friedrich-Schiller-Universität Jena

https://wwbp.org [1,2] https://julielab.de [3]

Abstract

Computational detection and understanding of empathy is an important factor in advancing human-computer interaction. Yet to date, text-based empathy prediction has the following major limitations: It underestimates the psychological complexity of the phenomenon, adheres to a weak notion of ground truth where empathic states are ascribed by third parties, and lacks a shared corpus. In contrast, this contribution presents the first publicly available gold standard for empathy prediction. It is constructed using a novel annotation methodology which reliably captures empathy assessments by the writer of a statement using multi-item scales. This is also the first computational work distinguishing between multiple forms of empathy, empathic concern, and personal distress, as recognized throughout psychology. Finally, we present experimental results for three different predictive models, of which a CNN performs the best.

1 Introduction

Over two decades after the seminal work by Picard (1997) the quest of *Affective Computing*, to ease the interaction with computers by giving them a sense of how emotions shape our perception and behavior, is still far from being fulfilled. Undoubtedly, major progress has been made in NLP, with sentiment analysis being one of the most vivid and productive areas in recent years (Liu, 2015).

However, the vast majority of contributions has focused on *polarity prediction*, typically only distinguishing between positive and negative feeling or evaluation, usually in social media postings or product reviews (Rosenthal et al., 2017; Socher et al., 2013). Only very recently, researchers started exploring more sophisticated models of human emotion on a larger scale (Wang et al., 2016; Abdul-Mageed and Ungar, 2017; Mohammad and Bravo-Marquez, 2017a; Buechel and Hahn, 2017, 2018a,b). Yet such approaches, often rooted in psychological theory, also turned out to be more challenging in respect to annotation and modeling (Strapparava and Mihalcea, 2007).

Surprisingly, one of the most valuable affective phenomena for improving human-machine interaction has received surprisingly little attention: *Empathy*. Prior work focused mostly on *spoken dialogue*, commonly addressing conversational agents, psychological interventions, or call center applications (McQuiggan and Lester, 2007; Fung et al., 2016; Pérez-Rosas et al., 2017; Alam et al., 2017).

In contrast, to the best of our knowledge, only three contributions (Xiao et al., 2012; Gibson et al., 2015; Khanpour et al., 2017) previously addressed *text-based* empathy prediction[1] (see Section 4 for details). Yet, all of them are limited in three ways: (a) neither of their corpora are available leaving the NLP community without shared data, (b) empathy ratings were provided by others than the one actually experiencing it which qualifies only as a weak form of ground truth, and (c) their notion of empathy is quite basic, falling short of current and past theory.

* These authors contributed equally to this work. Anneke Buffone designed and supervised the crowdsourcing task and the survey described in Section 2, and provided psychological background knowledge. Sven Buechel was responsible for corpus creation, data analysis, and modeling. The technical set-up of the crowdsourcing task and the survey was done jointly by both first authors.

†Work conducted while being at the University of Pennsylvania.

[1] Psychological studies commonly distinguish between *state* and *trait* empathy. While the former construct describes the amount of empathy a person experiences as a direct result of encountering a given stimulus, the latter refers to how empathetic one is on average and across situations. This studies exclusively addresses *state empathy*. For a contribution addressing *trait empathy* from an NLP perspective, see Abdul-Mageed et al. (2017).

Authors: Sven Buechel, Anneke Buffone, Barry Slaff, Lyle Ungar, Joao Sedoc
Venue: EMNLP
Year: 2018

Definition of Empathy-related Terms

- Drawn from psychology - specifically the work of Batson et al. (1987)
- **Empathic Concern** - Warm, compassionate feeling towards a suffering target
  - Maintains a separation between the self and the other entity, and the feeling of empathy itself is focused on the other entity rather than the self
- **Personal Distress** - Self-focused, negative affect state that occurs as the result of witnessing another entity's suffering or unfulfilled needs
- The main distinctions between empathic concern and personal distress are:
  - Sentiment
  - Focus
- The authors acknowledge that there are many psychologists who would use a more expanded definition for these terms, but these are the operational definitions being used for this study specifically

Data Collection

- Two research interns collected a total of 418 articles likely to evoke empathic reactions in readers
- The researchers then set up a crowdsourcing task on Amazon Mechanical Turk in which participants read a random selection of five news articles
- After each article, they were rated on their empathic concern and personal distress levels (more on this in the next slide), and they were instructed to write a short message about their feelings and thoughts regarding the article

> ing as follows: *Now that you have read this article, please write a message to a friend or friends about your feelings and thoughts regarding the article you just read. This could be a private message to a friend or something you would post on social media. Please do not identify your intended friend(s) — just write your thoughts about the article as if you were communicating with them. Please use between 300 and 800 characters.*

Data Annotation

- In order to measure empathic concern and personal distress, the researchers used Batson's Empathic Concern – Personal Distress Scale (Batson et al., 1987)
- After reading the articles and before writing their message to a friend, participants ranked how they felt various emotions on a scale of 1-7
- Each emotion corresponded either to empathic concern or personal distress
- These emotions were averaged out to create ratings for empathic concern and personal distress

How strongly do you feel the following emotion? Using the 1-7 scale below, please indicate your agreement.

	Not at All 1	2	3	4	5	6	Extremely 7
Warm	O	O	O	O	O	O	O
Tender	O	O	O	O	O	O	O
Sympathetic	O	O	O	O	O	O	O
Softhearted	O	O	O	O	O	O	O
Moved	O	O	O	O	O	O	O
Compassionate	O	O	O	O	O	O	O
Worried	O	O	O	O	O	O	O
Upset	O	O	O	O	O	O	O
Troubled	O	O	O	O	O	O	O
Perturbed	O	O	O	O	O	O	O
Grieved	O	O	O	O	O	O	O
Disturbed	O	O	O	O	O	O	O
Alarmed	O	O	O	O	O	O	O
Distressed	O	O	O	O	O	O	O

Figure 2: Multi-item scales for empathic concern and personal distress.

Data Reliability

- Inter-rater reliability does not make sense for this type of data.
- Instead, the researchers used split-half reliability, which is calculated using the following process:
  - Split the participants’ ratings for each individual word (*warm*, *tender*, etc.) randomly into two groups.
  - Average the individual item ratings for each group.
  - Measure the correlation between the two groups for each item.
  - The process is repeated 100 times with random splits, and the correlations are averaged to get the final SHR.
- Empathy: $r = 0.875$
- Distress: $r = 0.924$

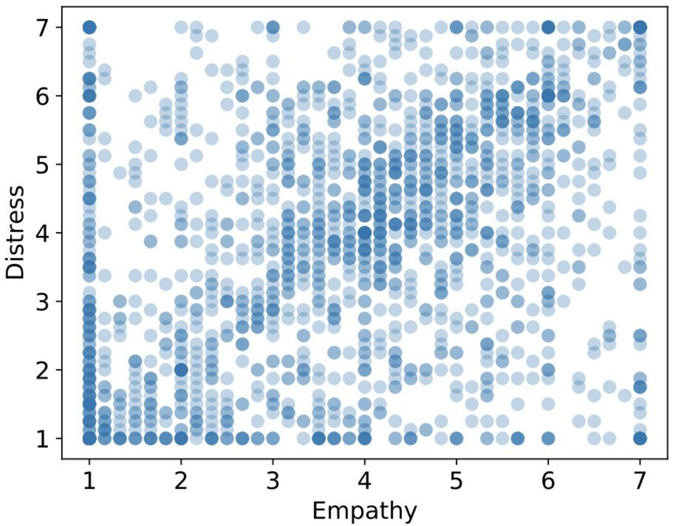
Corpus Structure

	E	D	Message
(1)	4.8	3.1	*I'm sorry to hear that about Dakota's parents. Even when you are adult it must be hard to see your parents splitting up. No one wants that to happen and it's unfortunate that her parents couldn't work it out. I hope they are able to still remain civil around the kids and family. Just because it didn't work romantically doesn't mean it won't work at all.*
(2)	4.0	5.5	*Here's an article about crazed person who murdered two unfortunate women overseas. Life is crazy. I can't imagine what the families are going through. Having to go to or being forced into sex work is bad enough, but for it to end like this is just sad. It feels like there's no place safe in this world to be a woman sometimes.*
(3)	1.0	1.3	*I just read an article about some chowder-head who used a hammer and a pick ax to destroy Donald Trump's star on the Hollywood walk of fame. Wow, what a great protest. You sure showed him. Good job. Lol, can you believe this garbage? Who has such a hollow and pathetic life that they don't have anything better to do with their time than commit petty vandalism because they dislike some politician? What a dingus.*

Table 1: Illustrative examples from our newly created gold standard with ratings for empathy (**E**) and distress (**D**).

Corpus Analysis

- Direct linear correlation between empathy and distress, exceptions are either:
  - Extremely low-empathy
  - Or extremely low-distress
- What this shows us is that empathy and distress are distinct phenomenon, in that you can have one without the other, but in many cases, they show up together because they result from the same type of stimulus



Regression Models

- Regression models were trained on the corpus and tested using 10-fold cross validation
  - Ridge
  - Feedforward Neural Network (FNN)
  - Convolutional Neural Network (CNN)
- The best-performing model for both empathy and distress was the **CNN**; however, it only achieved a mean correlation coefficient of r = 0.424
- The researchers note that this correlation coefficient is very similar to those found in the early days of emotion detection and predict that this correlation coefficient could double with a larger corpus size and more research just as it did with emotion detection

	Empathy	Distress	Mean
Ridge	.385	.410	.398
FFN	.379	.401	.390
CNN	**.404***	**.444***	**.424***

Table 2: Model performance for predicting empathy and distress in Pearson's r; with row-wise mean; best result per column in bold, significant ($p < .05$) improvement over other models marked with '*'.

Discussion

- What do you think about the distinction between empathic concern and personal distress?
- What type of model would you use to predict empathic concern and personal distress from text?

[Additional Paper] WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories

- **Shared Task**: Multiple teams compete to design the best model for some dataset
  - Each team is encouraged to submit a separate paper detailing their highest performing implementation
- This shared task aims to achieve the highest performance on an extended version of the empathy and distress dataset created in the previous paper
  - Task 1: Predict empathy/distress
  - Task 2: Predict emotion labels

WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories

Shabnam Tafreshi
Georgetown University
st1093@georgetown.edu

Orphée De Clercq
LT3, Ghent University
orphee.declercq@ugent.be

Valentin Barriere
European Commission Joint Research Centre
Valentin.BARRIERE@ec.europa.eu

João Sedoc
New York University
jsedoc@stern.nyu.edu

Sven Buechel
Friedrich Schiller University Jena
sven.buechel@uni-jena.de

Alexandra Balahur
European Commission Joint Research Centre
alexandra.balahur@ec.europa.eu

Abstract

This paper presents the results that were obtained from the WASSA 2021 shared task on predicting empathy and emotions. The participants were given access to a dataset comprising empathic reactions to news stories where harm is done to a person, group, or other. These reactions consist of essays, Batson empathic concern, and personal distress scores, and the dataset was further extended with news articles, person-level demographic information (age, gender, ethnicity, income, education level), and personality information. Additionally, emotion labels, namely Ekman's six basic emotions, were added to the essays at both the document and sentence level. Participation was encouraged in two tracks: predicting empathy and predicting emotion categories. In total five teams participated in the shared task. We summarize the methods and resources used by the participating teams.

1 Introduction

It is important to be able to analyze empathy and emotion in natural languages. Emotion classification in natural languages has been studied over two decades and many applications successfully used emotion as their major components. Empathy utterances can be emotional, therefore, examining emotion in text-based empathy possibly has a major impact on predicting empathy. Analyzing text-based empathy and emotion have different applications; empathy is a crucial component in applications such as empathic AI agents, effective gesturing of robots, and mental health, emotion has natural language applications such as commerce, public health, and disaster management. In this paper, we present the WASSA 2021 Shared Task: Predicting Empathy and Emotion in Reaction to News Stories. This shared task included two individual tasks where teams develop models to predict emotions and empathy in essays in which people expressed their empathy and distress in reaction to news articles in which an individual, group of people or nature was harmed. Additionally, the dataset also included the demographic information of the authors of the essays such as age, gender, ethnicity, income, and education level, and personality information (details of the collection of the dataset is provided in section 3). Optionally, we suggested that the teams could also use emotion labels when modeling empathy to learn more about the impact of emotions on empathy. The shared task consisted of two tracks:

1. Predicting Empathy (EMP): the formulation of this track is to predict the Batson empathic concern ("feeling for someone") and personal distress ("suffering with someone") using the essay, personality information, demographic information, and emotion.
2. Emotion Label Prediction (EMO): the formulation of this track is to predict emotion tags

92

Proceedings of the 11th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 92–104
April 19, 2021. ©2021 Association for Computational Linguistics

Authors: Shabnam Tafreshi, Orphée De Clercq, Valentin Barriere, João Sedoc, Sven Buechel, Alexandra Balahur
Venue: WASSA (Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis)
Year: 2021

Dataset Extension

- The MTurk workers in the original study filled out demographic surveys and personality tests; however, this information was not included in the original corpus
- The researchers in this study encode these attributes into the dataset and provide it to the shared task participants as optional inputs for their model
- Demographic information includes:
  - Age
  - Gender
  - Ethnicity
  - Income
  - Education Level
- Personality is represented by:
  - Big Five, also known as OCEAN (Gosling et al., 2003)
  - Interpersonal Reactivity Index (Davis, 1980)
  - **Note**: The Satisfaction With Life Scale (Diener et al., 1985) was collected in the original study but not used in either dataset
- Authors added labels identifying which one of Ekman's six basic emotions each essay in the corpus was expressing
  - Did not specify how this was done (manual, crowdsourcing, etc.)

Machine Learning Architectures

- Every system utilized supervised machine learning models for both tasks
- Most teams used pretrained transformer language models and either:
  - Fine-tuned them using the training data
  - Or extracted features from different layers to use in their final model
- **IITK** - Indian Institute of Technology, Kanpour (Mundra et al., 2021)
  - Fine-tuned the ELECTRA-Large transformer model with a single FFN on top of it
  - Fine-tuned a RoBERTa transformer model using multi-task learning
  - Final model was an ensemble of these two
- **PVG** - Pune University (Kulkarni et al., 2021)
  - Multi-input, multi-task transformer architecture
  - Uses empathy score prediction as its primary task and emotion classification as a secondary task
- **MilaNLP** - Quebec A.I. Institute a.k.a. MILA (Fornaciari et al., 2021)
  - Found that multi-task learning and using demographic/personality inputs led to worse performance than single-task learning
    - Suggests that emotion and empathy/distress are not related tasks
  - Used the BERT language model (Devlin et al., 2019) w/ single-task learning
- **Team Phoenix** - Indian Institute of Technology, Kharagpur (Butala et al., 2021)
  - Used a multi-layer perceptron network for the empathy/distress prediction task
  - Used Google's T5-Finetuned transformer for emotion recognition (Raffel et al., 2019) and fine-tuned it on the training data for the emotion classification task
- **EmpNa** - Centrica and the National Research Council of Italy (Vettigli et al., 2021)
  - Used linear regression on unigrams, bigrams, and trigrams for empathy/distress prediction
  - Used logistic regression on the same lexical features for emotion classification

Results

Team	Emp	Dis	Avg
PVG	0.517	**0.574**	**0.545**
EmpNa	0.516	0.554	0.536
WASSA@IITK	**0.558**	0.507	0.533
Team Phoenix	0.358	0.476	0.417

Table 3: Results of the teams participating in the EMP track (Pearson correlations).

Team	P	R	F1	Acc
WASSA@IITK	**0.57**	**0.55**	**0.55**	**0.62**
Team Phoenix	0.55	0.48	0.50	0.59
MilaNLP	0.55	0.47	0.49	0.58
EmpNa	0.32	0.31	0.31	0.40

Table 5: Results of the teams participating in the EMO track (macro-averaged precision (P), recall (R), F1-score (F1) and accuracy (Acc)).

	Joy			Sadness			Disgust			Fear			Anger			Surprise		
Team	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
WASSA@IITK	**40**	**53**	**46**	72	**55**	**63**	**70**	**42**	**52**	46	42	44	**74**	80	**77**	36	**50**	**42**
Team Phoenix	35	45	40	67	37	48	57	38	45	**52**	40	**45**	67	**83**	74	**48**	33	39
MilaNLP	34	38	36	**77**	31	44	58	38	46	34	48	40	73	81	76	**48**	33	39
EmpNa	23	28	25	31	25	28	28	30	29	30	28	29	60	59	59	10	15	12

Table 6: Breakdown EMO labels (MACRO)

Discussion

- Which model do you like the best and why?
- Are there any modeling approaches that you would try out for this task?

[Additional Paper] MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations

- Builds off **EmotionLines** ([Chen et al., 2018](#))
  - Dataset of conversations that labels each speaker utterance with an emotion
  - Unimodal - only contains text
  - Extracted from the T.V. series *Friends* and labeled using five crowdsourced workers Amazon MTurk
- Adds two new modes
  - Audio
  - Visual (Video)
- MELD: Multimodal EmotionLines Dataset

MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations

Soujanya Poria[†], Devamanyu Hazarika[Φ], Navonil Majumder[‡], Gautam Naik[¶], Erik Cambria[¶], Rada Mihalcea[ι]

[†]Information Systems Technology and Design, SUTD, Singapore
[Φ]School of Computing, National University of Singapore, Singapore
[‡]Centro de Investigación en Computación, Instituto Politécnico Nacional, Mexico
[¶]Computer Science & Engineering, Nanyang Technological University, Singapore
[ι]Computer Science & Engineering, University of Michigan, USA
sporia@sutd.edu.sg, hazarika@comp.nus.edu.sg, navo@nlp.cic.ipn.mx, gautam@sentic.net, cambria@ntu.edu.sg, mihalcea@umich.edu

arXiv:1810.02508v6 [cs.CL] 4 Jun 2019

Abstract

Emotion recognition in conversations (ERC) is a challenging task that has recently gained popularity due to its potential applications. Until now, however, there has been no large-scale multimodal multi-party emotional conversational database containing more than two speakers per dialogue. To address this gap, we propose the *Multimodal EmotionLines Dataset* (MELD), an extension and enhancement of EmotionLines. MELD contains about 13,000 utterances from 1,433 dialogues from the TV-series *Friends*. Each utterance is annotated with emotion and sentiment labels, and encompasses audio, visual, and textual modalities. We propose several strong multimodal baselines and show the importance of contextual and multimodal information for emotion recognition in conversations. The full dataset is available for use at http://affective-meld.github.io.

1 Introduction

With the rapid growth of Artificial Intelligence (AI), multimodal emotion recognition has become a major research topic, primarily due to its potential applications in many challenging tasks, such as dialogue generation, user behavior understanding, multimodal interaction, and others. A conversational emotion recognition system can be used to generate appropriate responses by analyzing user emotions (Zhou et al., 2017; Rashkin et al., 2018).

Although significant research work has been carried out on multimodal emotion recognition using audio, visual, and text modalities (Zadeh et al., 2016a; Wollmer et al., 2013), significantly less work has been devoted to emotion recognition in conversations (ERC). One main reason for this is the lack of a large multimodal conversational dataset.

According to Poria et al. (2019), ERC presents several challenges such as conversational context modeling, emotion shift of the interlocutors, and others, which make the task more difficult to address. Recent work proposes solutions based on multimodal memory networks (Hazarika et al., 2018). However, they are mostly limited to dyadic conversations, and thus not scalable to ERC with multiple interlocutors. This calls for a multi-party conversational data resource that can encourage research in this direction.

In a conversation, the participants' utterances generally depend on their conversational context. This is also true for their associated emotions. In other words, the context acts as a set of parameters that may influence a person to speak an utterance while expressing a certain emotion. Modeling this context can be done in different ways, e.g., by using recurrent neural networks (RNNs) and memory networks (Hazarika et al., 2018; Poria et al., 2017; Serban et al., 2017). Figure 1 shows an example where the speakers change their emotions (emotion shifts) as the dialogue develops. The emotional dynamics here depend on both the previous utterances and their associated emotions. For example, the emotion shift in utterance eight (in the figure) is hard to determine unless cues are taken from the facial expressions and the conversational history of both speakers. Modeling such complex inter-speaker dependencies is one of the major challenges in conversational modeling.

Conversation in its natural form is multimodal. In dialogues, we rely on others' facial expressions, vocal tonality, language, and gestures to anticipate their stance. For emotion recognition, multimodal-

Authors: Soujanya Poria, Devmanyu Hazarika, Navonil Majunder, Gautam Naik, Erik Cambria, Rada Mihalcea
Venue: ACL
Year: 2019

Importance of Multimodality in Emotion Recognition

- Oftentimes, raw text does not give us enough information to adequately assess emotion, especially in dialogue
  - Text may be ambiguous
  - Or it may be said sarcastically in a way that can only be picked up through facial/vocal cues

Utterance: *"Become a drama critic!"*

Emotion: *Joy* **Sentiment:** *Positive*

Text	Audio	Visual
Ambiguous	Joyous tone	Smiling Face

Utterance: *"Great, now he is waving back"*

Emotion: *Disgust* **Sentiment:** *Negative*

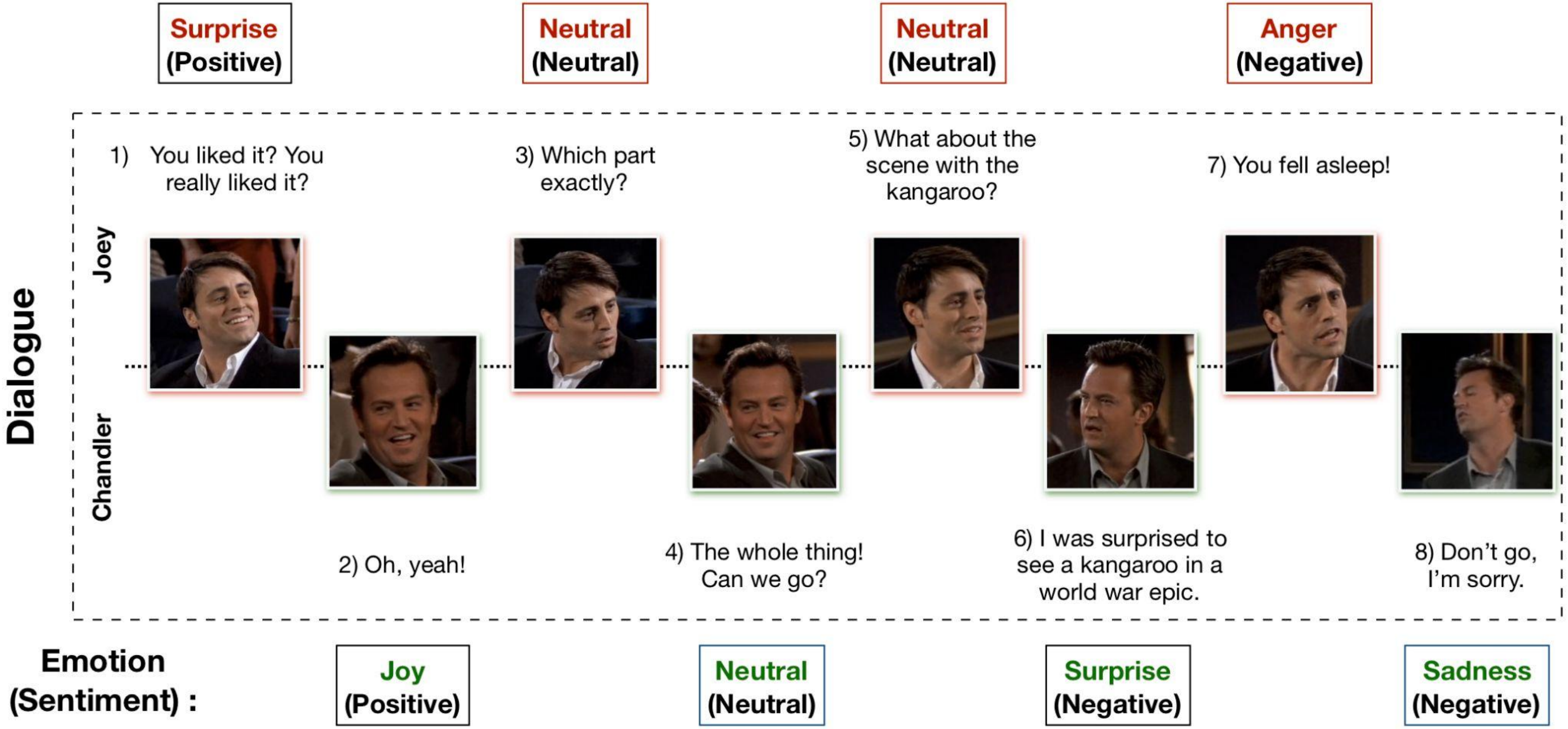
Text	Audio	Visual
Positive/Joy	Flat tone	Frown

Figure 2: Importance of multimodal cues. Green shows primary modalities responsible for sentiment and emotion.

Methodology for MELD

- The researchers extracted the timestamps corresponding to each utterance in the EmotionLines dataset using a timestamp alignment tool called Gentle
- They used these timestamps to extract the corresponding clip and had three graduate students "re-annotate" the utterances using those clips
  - Some utterances had their corresponding emotions changed, now that there was multimodal information for the annotators to evaluate
  - In the case where all three annotations were different for an utterance, the entire dialogue corresponding to it was dropped
  - Otherwise, disagreements were resolved via a majority voting system
  - **Fleiss's kappa**: 0.34
    - Characterized as "fair agreement" by Landis and Koch (1977)

MELD: Example Dialogue



Discussion Questions

- What do you think of the use of the T.V. sitcom *Friends* for building an emotion corpus? Potential concerns include:
  - Actors are trained to express their emotions in a much more exaggerated fashion than everyday people, both through their voice and their body language
  - *Friends* relies a lot on the laugh track and “awwws” between utterances in order to express the emotional state of the characters (see this clip of *Friends* without the laugh track)
  - Sitcom characters in general are often flanderized (or turned into caricatures of themselves) as the show progresses to later seasons, meaning that the emotions they express will be more and more likely to spring from their exaggerated personality traits rather than everyday things that would cause emotional reactions in the average person

Ethics of Emotion Detection Research

- In 2019, the AI Now Institute at NYU, in its annual report on the social implications of AI technologies, called for a ban on the use of affect recognition technology in certain cases, specifically:
  - Employment
    - ex. HireVue, an automated interviewing platform that claims to be able to predict a candidate's "grit" and tracks how often the candidate smiles, among other things
  - Education
    - ex. BrainCo, a Boston-based company, is creating AI headbands that they claim can monitor students' brain activity and send them to teachers and parents via a mobile app to verify whether or not they're paying attention in school
  - Criminal Justice
    - ex. EyeDetect, a system that flags potential deception via eye movements and pupil dilation, is used by local and federal government to screen job applicants and assess the veracity of criminal suspects' claims

Ethics of Emotion Detection Research (cont.)

- The main concerns with emotion recognition technology (especially those based on the visual or audio modalities) is that they can:
  - Amplify biases, especially towards marginalized groups
    - For audio modalities, this could be via an accent or dialect (such as AAVE)
    - For visual modalities, emotion recognition systems may not be as effective on darker skin tones (when all the training data uses light-skinned models) and often characterize the facial features of certain ethnic groups as corresponding to certain emotions
    - Rhue (2018) found that, when running face-based emotion detection systems on NBA players, African Americans were assigned more negative emotion scores than Whites, regardless of how much they smiled
  - Be inaccurate and harm those that are misassessed by them
    - As discussed in the previous slide, emotion recognition technologies are being used to deny people access to opportunities
  - Judges people on the basis of physical characteristics that they don't fully have control over
    - Many people have resting facial expressions that don't necessarily correspond to the emotion they are currently feeling

Discussion

- What factors should be considered when assessing the ethics of a specific emotion detection system?
- Does the ethicality of affect recognition technology depend on the modality it uses?
    - Is it possible that one modality (text, for example) carries less biases than others?

Online Resources

- Demos
  - DeepMoji
  - IBM Watson Natural Language Understanding (NLU)
  - IBM Watson Tone Analyzer
  - IBM Watson Tone Analyzer for Customer Engagement
- Libraries
  - IBM Watson NLU Emotion Detection API
- Repositories
  - IMS EmoInt
  - DeepMoji

Discussion

- What is the best way to collect reliable, labeled data for emotion detection research?
- How useful is this research to society? Do the pros outweigh the cons?
- If you had $1,000,000 to invest in emotion detection research, how would you split it between: psychologists, neuroscientists, computer scientists, and philosophers?
- Any other thoughts?