Understanding the Role of Affect Dimensions in Detecting Emotions from Tweets: A Multi-task Approach

Rajdeep Mukherjee[1], Atharva Naik[1], Sriyash Poddar[1], Soham Dasgupta[2], Niloy Ganguly[1]

[1]CNeRG Lab, Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur
[2]Mallya Aditi International School, Bangalore, India

sigir21

Psychological Models of Emotions

Categorical Models: classify affective states into discrete categories. Paul Ekman, for example, proposed the existence of six basic, distinct and universal emotions.

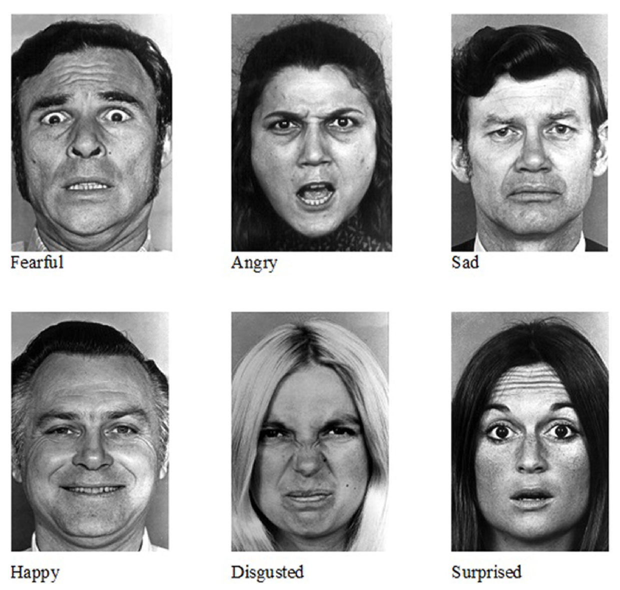

Figure: Ekman's 6 Basic Emotions

Dimensional Models: propose fundamental affect dimensions that constitute emotional spaces. Russell's VAD model, for e.g., interprets emotions as points in a 3-D space.

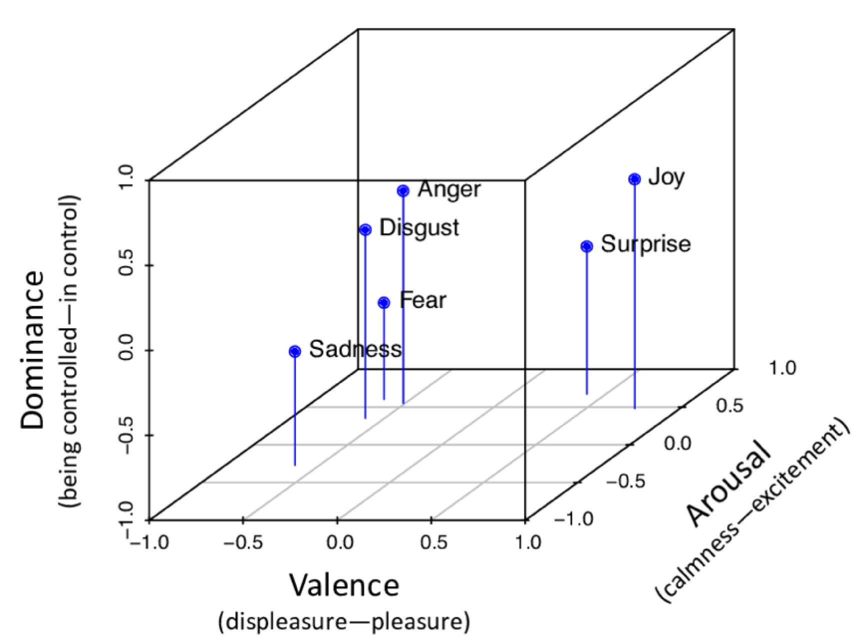

Figure: Russell's Valence-Arousal-Dominance Model

Relationship? Emotional words (excited, for e.g.) and sentence context (win vs What a win!) carry scores along multiple dimensions - help in capturing the overall emotion.

I was ***so*** ***excited*** to be a part of this team! What a ***win***!

- *Ekman Emotion*: Joy
- *Valence*: 0.84, *Arousal*: 0.84, *Dominance*: 0.88
- *so* - high Arousal score, *excited* - high Valence, and high Arousal scores.
- *win* - high Valence, and high Dominance scores.

I was so excited to be a part of this team! But, alas we ***lost***.

- *Ekman Emotion*: Neutral
- *Valence*: 0.58, *Arousal*: 0.84, *Dominance*: 0.36
- *lost* - low Valence, and low Dominance scores.

The Objectives

- Build a system that efficiently detects emotions from written narratives, especially tweets.
- To investigate if VAD supervisions can improve the emotion classification (EC) performance.

The Task

Figure: Given a tweet, identify the emotion(s) expressed in it.

VADEC Architecture

We propose VADEC, that co-trains multi-label emotion classification (*classifier module*) and multi-dimensional emotion regression (*regressor*) by jointly optimizing the weights of a shared text-encoder, based on BERTweet.

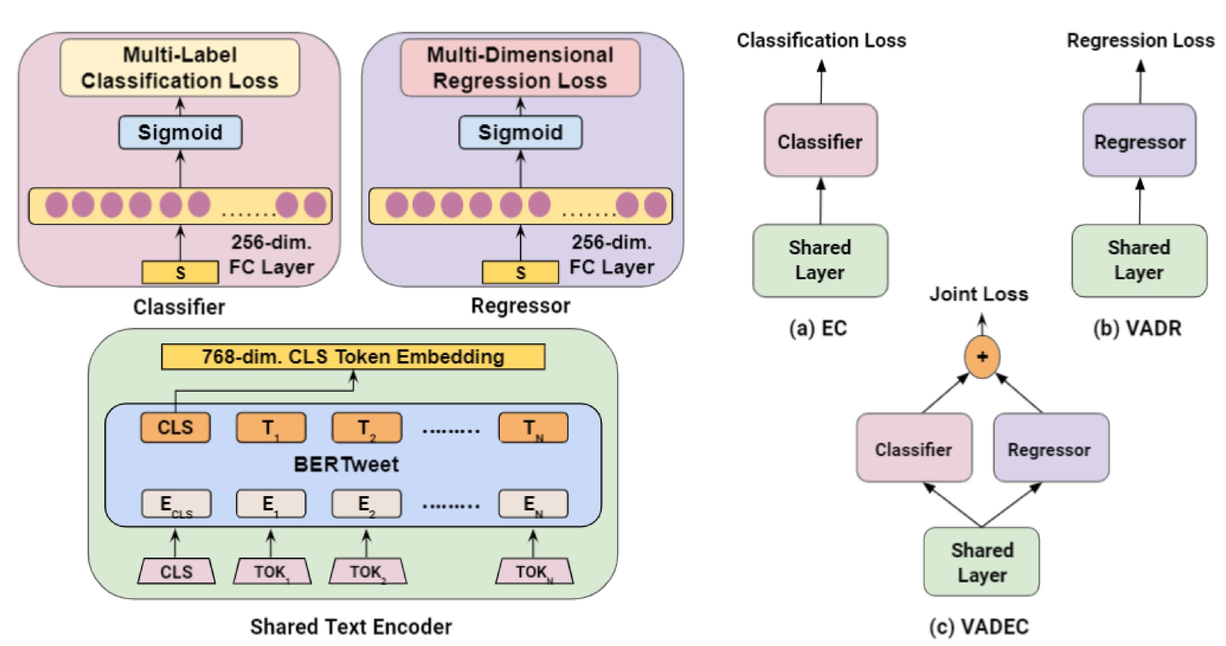

Figure: Model Components and Architecture

Datasets

- ***EmoBank***: VAD dataset - around 10K sentences annotated with continuous scores for V, A, and D.
- ***AIT***: - EC dataset - 10,983 English tweets annotated for the presence/absence of 11 general emotions.
- ***SenWave***: EC dataset - 10K tweets annotated for the presence/absence of 11 emotions specific to COVID-19.

Experiments and Results

Emotion Classification Models:

- *BERTL:* BERT-Large fine-tuned on the *AIT* dataset.
- *NTUA-SLP:* Bi-LSTM with multi-layer self attentions, pre-trained on a large collection of general tweets before fine-tuning the model on the *AIT*. This further represents the winning entry in SemEval 2018 Task 1.
- *SenWave:* Scores reported by the authors of SenWave.
- *EC:* Our classifier module, when trained as a single task.
- *$EC_{RoBERTa}$:* *EC* with RoBERTa as the shared layer, instead of BERTweet, a model ablation.

Emotion Regression Models:

- *AAN:* uses adversarial learning between two attention layers to learn discriminative word-weight parameters for scoring two emotion dimensions at a time.
- *All_In_One:* uses a multi-task ensemble framework to learn different configurations of tasks related to coarse- and fine-grained sentiment and emotion analysis.
- *SVR-SLSTM:* A semi-supervised approach using variational autoencoders to predict the VAD scores.
- *BERTL (EB ← AIT):* BERT-Large, first fine-tuned on *AIT*, followed by fine-tuning on *EmoBank*.
- *VADR:* Our regressor module, trained as a single task.
- *$VADR_{RoBERTa}$:* *VADR* with RoBERTa as the shared layer, instead of BERTweet, a model ablation.
- *VADEC (AIT), and VADEC (SenWave):* VADEC trained respectively with *AIT*, and *SenWave*.

Methods	Jaccard Acc.	F1-Macro	F1-Micro
BERTL	0.572	0.534	0.697
NTUA-SLP	0.588	0.528	0.701
$EC_{RoBERTa}$	0.592	0.570	0.712
EC	0.605	0.581	0.723
VADEC	**0.608**	**0.593**	**0.728**
Significance T-Test (p-values)	0.029	-	-

Figure: Multi-label Emotion Classification Results on the *AIT*.

Methods	Accuracy	Jac. Acc.	F1-Macro	F1-Micro	LRAP	Ham. Loss
SenWave	0.847	0.495	0.517	0.573	0.745	0.153
VADEC	**0.877**	**0.560**	**0.563**	**0.620**	**0.818**	**0.123**

Figure: Multi-label Emotion Classification Results on the *SenWave*.

Methods	Valence (V)	Arousal (A)	Dominance (D)
AAN	0.424	0.351	0.265
All_In_One	0.635	0.375	0.277
SRV-SLSTM	0.620	0.508	0.333
BERTL (EB ← AIT)	0.765	**0.583**	0.416
$VADR_{RoBERTa}$	0.804	0.494	**0.511**
VADR	0.821	0.553	0.493
VADEC (AIT)	0.820	0.563	0.459
VADEC (SenWave)	**0.823**	0.553	0.485

Figure: Comparison of Pearson Correlation (r-values) for the Multi-dimensional Emotion Regression task on the *EmoBank* dataset.

COVID-19 and Indians: A Case Study

We use VADEC, trained on *EmoBank* and *SenWave* to detect and analyze the changing dynamics of Indian emotions towards the COVID-19 pandemic from their tweets.

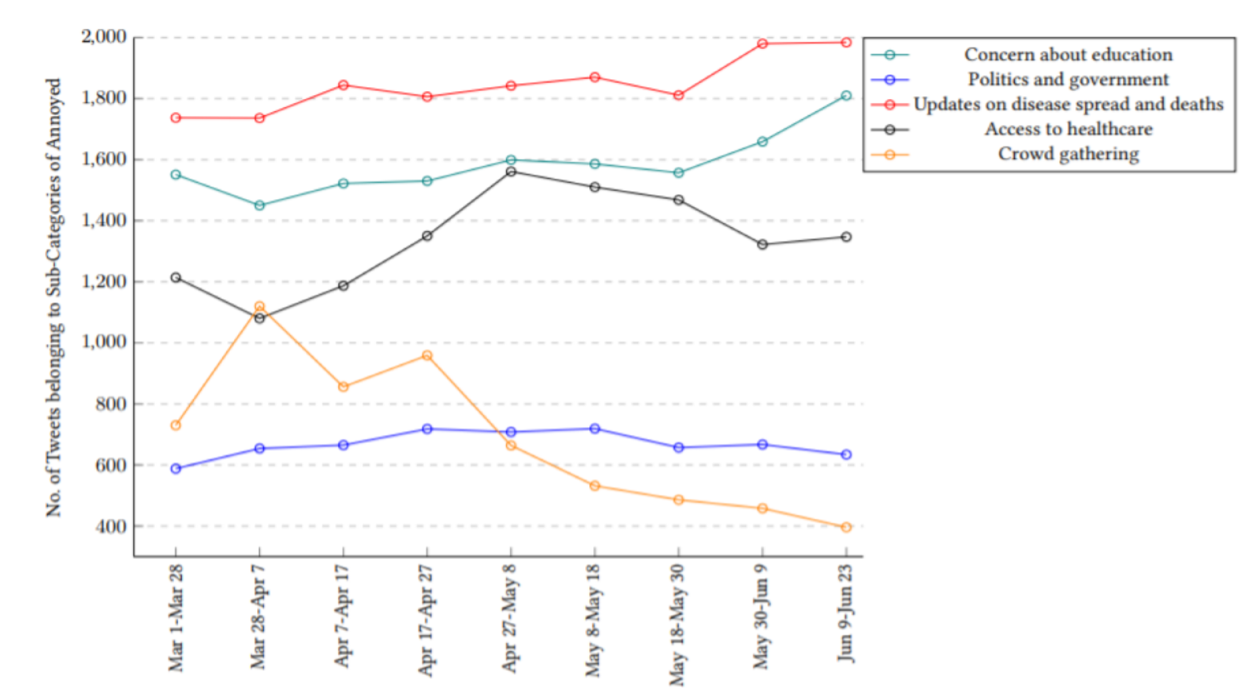

Figure: Change in Sub-categories of Emotional Triggers for *Annoyed* towards the COVID-19 pandemic over time.

Detailed results are available in the paper.

Key Contributions

We show that the performance of emotion recognition in written narratives, especially tweets, can be improved by utilizing the better representational power of the dimensional models of emotion representation.

For Further Information

Preprint: `https://arxiv.org/abs/2105.03983`
Open-source Implementation: `https://github.com/atharva-naik/VADEC`