

Tweet Emotion Dynamics: Emotion Word Usage in Tweets from US and Canada

Krishnapriya Vishnubhotla (vkpriya@cs.toronto.edu)
Saif M. Mohammad (saif.mohammad@nrc-cnrc.gc.ca)

Department of Computer Science, University of Toronto, Vector Institute for Artificial Intelligence
National Research Council Canada

We introduce:

- TUSC:** ~45 million geo-located Tweets from US, Canada
- Tweet Emotion Dynamics (TED):** a set of metrics to quantify emotion word usage in tweets across time

Why Track Emotion Word Usage

- Over Time:** Are we tweeting more positive words, negative words, high arousal words, etc. over time?
How has the COVID-19 pandemic impacted tweets?
When did we use the most number of words conveying a lack of control and uncertainty?
How were individual cities impacted? etc.
- Geographically:** How are Canada and US different in terms of emotion word usage?

The TUSC Dataset (after appropriate preprocessing of tweets)

- TUSC-Country:** tweets from US and Canada (CA)
 - ~103K tweets per year, 2015–2018
 - ~380K tweets per year, 2019–2021
- TUSC-City:** tweets from 46 US–CA cities, Apr 2020–Dec 2021
 - ~26M tweets per year

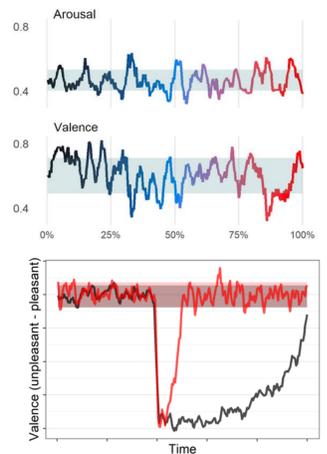
Fun fact: CA tweets, on average, use two more tokens per tweet than US tweets (less informal, slang, etc. (Sneffjella et al. 2018))

Utterance (e.g., Tweet) Emotion Dynamics

- Emotion Dynamics:** A framework from Psychology for measuring changes in one’s emotions over time.
- Utterance Emotion Dynamics (UED) Hipson and Mohammad (2020):** A computational framework that captures changes in emotions associated with *utterances* over time.
- Tweet Emotion Dynamics (TED):** Our use of UED on tweets.

Set Up

- Create temporally ordered stream of words by speaker
- Apply a rolling window averaging **word–emotion association scores**
- This sequence of window emotion scores is the **emotion arc** of their utterances



Emotion dimensions

- Valence (V):** positive–negative
pleasure–displeasure
- Arousal (A):** active–sluggish
- Dominance (D):** powerful–weak
in control–no control

NRC VAD Lexicon

- Scores between 0 and 1 for the **V, A, D** for ~20,000 English words
- Removed ‘neutralish’ words (0.33–0.66)
- Removed frequent ambiguous words (*trump, may, will*)

Notable UED Metrics

- Home Base:** Range of emotion scores one standard deviation away from the mean on each side. Most probable emotion space occupied by speaker.
- Variability:** Standard deviation of emotion means.
- Rise Rate:** Rate at which speaker reaches peak emotional intensity (emotional reactivity).
- Recovery Rate:** Rate at which speaker returns to home base (emotion regulation).

EXPERIMENTS

1. Average Emotion Scores of Words in Tweets (How emotional are our tweets?)

Year		Valence		Arousal		Dominance	
		Canada	USA	Canada	USA	Canada	USA
2015	TUSC	0.675	0.644	0.461	0.465	0.564	0.535
2016	Country	0.672	0.644	0.464	0.469	0.567	0.543
2017		0.666	0.639	0.465	0.472	0.568	0.546
2018		0.669	0.640	0.463	0.472	0.576	0.553
2019		0.668	0.638	0.461	0.476	0.577	0.553
2015–2019		0.670	0.641	0.463	0.471	0.570	0.546
2020		0.662	0.634	0.458	0.474	0.574	0.555
2021		0.669	0.644	0.457	0.474	0.578	0.560
2020	TUSC	0.651	0.634	0.469	0.479	0.571	0.558
2021	City	0.658	0.645	0.470	0.479	0.576	0.563

CA (vs US): higher V, lower A, higher D

Yearly:

- V:** Lowest in 2020 (June 2020)
- A:** US scores increased with time
- D:** Steady increase with time

Pandemic impact: lower V, lower D

TUSC-City, Country: similar trends

2. Proportion of Tweets with Emotional Terms (How often are we tweeting emotional terms?)

Find proportion of tweets with at least one high **V/A/D** word

Divided lexicons into low (score ≤ 0.33) and high (score ≥ 0.67) sub-lexicons.

Year	Low Valence		High Valence		Low Arousal		High Arousal	
	Canada	USA	Canada	USA	Canada	USA	Canada	USA
2015	36.3	38.4	80.2	75.4	55.8	51.1	40.9	38.1
2016	37.6	39.6	80.8	76.7	56.4	52.0	42.1	40.0
2017	40.8	43.2	82.9	79.0	59.4	55.6	45.0	43.5
2018	42.7	45.7	83.2	80.2	61.4	57.8	47.1	46.0
2019	42.9	45.3	82.7	79.4	60.8	56.7	46.5	46.1
2015–2019	40.0	42.4	82.0	78.1	58.8	54.6	44.3	42.7
2020	43.8	46.1	82.3	79.2	60.8	57.0	45.7	45.9
2021	42.6	44.5	82.5	79.7	61.1	57.2	45.6	46.1

High-V (vs Low-V): ~ 100% more

Low-A (vs High-A): ~ 40% more

High-D (vs Low-D): ~33% more

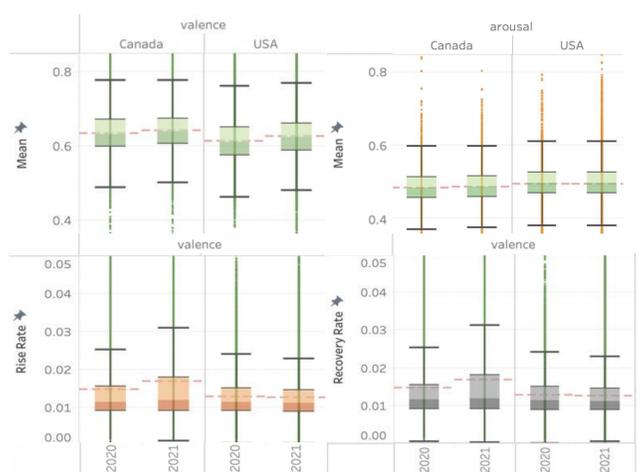
CA (vs US): more high-V, fewer low-V
more low-A, high-D

2020: higher low-V, fewer high-V

- We tweeted fewer positive **AND** more negative words

3. UED Metrics

Exps 1,2: aggregate emotion word usage (city, country)
Here we benchmark **individual** tweeter behavior.



- Overall:** Metrics show Gaussian distribution
- Mean:** Similar VAD trends across CA–US as in 1, 2
- Rise and Recovery Rates:** Larger third quartile for CA
V and D: rise and recovery rates for positive deviations from home base are lower than for negative deviations.
 - Trend is reversed for A
- City-level** (figure not shown here):
 - Highest V: London, Ottawa, Halifax, Victoria (CA)
 - Lowest V: Detroit, Houston, LA, Philadelphia (US)

Ethical Considerations

Before you start:

- Ethics Sheet for Automatic Emotion Recognition (Mohammad, 2022)
- Practical and Ethical Considerations in the Use of Emotion Lexicons (Mohammad, 2020)

Where Do We Go From Here?

- TED framework can be applied to any set of temporally-ordered utterances
- Applications in domains such as digital humanities, social sciences, psychology.
- Collaboration with UNC Carolina Affective Science Lab:
 - Can TED inform us about the mental and physical health of populations?