Paper Review #12

[CHI 2019] Moments of Change: Analyzing Peer-Based Cognitive Support In Online Mental Health Forum

Dept. of Computer Science & Engineering
202122029 Meeyun Kim

2022. 02. 17.

Moment of Change

- A **positive change in sentiment** for the OP* on a topic that was mentioned by the OP in their first post, over the course of a conversation in a single forum thread.

(***OP**: Original poster)

Deriving a Ground Truth

- Collected a sample of **2,500 posts** from **Talklife**.

- Given post on 7 point Likert scale (-3: strongly negative, 3: strongly positive)

- This dataset is called **annotation-based dataset.**

Deriving a Ground Truth (Cont'd)

- To create a larger scale dataset, qualitative analysis is conducted.

- **Regular expression-based phrases** are used to detect OP's sentiment.

- This dataset is called **pattern-based dataset.**

- **Gradient boosting classifier** is trained on the pattern-based labels, and tested on the crowdsourced labels.

Culture-specific datasets

- **Indians** vs **non-Indians**

- **Indians**: 25,537 threads without a moment of change,
  295 threads with a moment of change.

- **non-Indians**: 14,604 threads without a moment of change,
  6,396 threads with a moment of change.

Descriptive Analysis using Metadata

- Threads with moments of change have a **higher amount of interaction**.
  (9 message, 17 words < 12 messages, 27 words)

- Moments of change would happen in **OP's 7th response** (on average).

- Threads with moments of change are likely to have a **higher number of responders from the same country as the OP.**

Features for Predictive Analysis

(1) **LIWC-based**: positive, negative sentiment words, linguistic style matching

(2) **Punctuation-based Features**: the number of “?”, “!” …

(3) **Metadata-based Features**: number of posts in a thread, length of post,
number of countries …

(4) **Mental Health Language-based Features**:
250 most popular tri/four-grams from the Anxiety, Depression, Suicide Watch Reddit communities in 2015, names of medication from the Wikipedia article.

Results From Predictive Models

- Train : Validation : Test = 6,410 : 713 : 791

- Using **XGBoost**.

Results From Predictive Models (Cont'd)

Thread-level AUC

	LIWC	LIWC + Punctuation	LIWC + Punctuation + Metadata	LIWC + Punctuation + Metadata + Language
CA Dataset	0.87	0.87	0.88	0.88
CA Dataset, only non-OP posts	0.86	0.85	0.85	0.86
CA Dataset, only OP posts	0.68	0.69	0.81	0.81
Indian Dataset	0.89	0.9	0.9	0.9
Indian Dataset, only non-OP posts	0.97	0.97	0.98	0.98
Indian Dataset, only OP posts	0.78	0.73	0.92	0.93

Table 1: Thread-level AUC for models trained on the Culturally Agnostic and the Indian dataset. For both, we obtain high AUC scores (> 0.8) for predicting a moment of change. Models trained on non-OP LIWC features perform better than those trained on OP-only features, suggesting that responders' language plays an important role in detecting a moment of change.

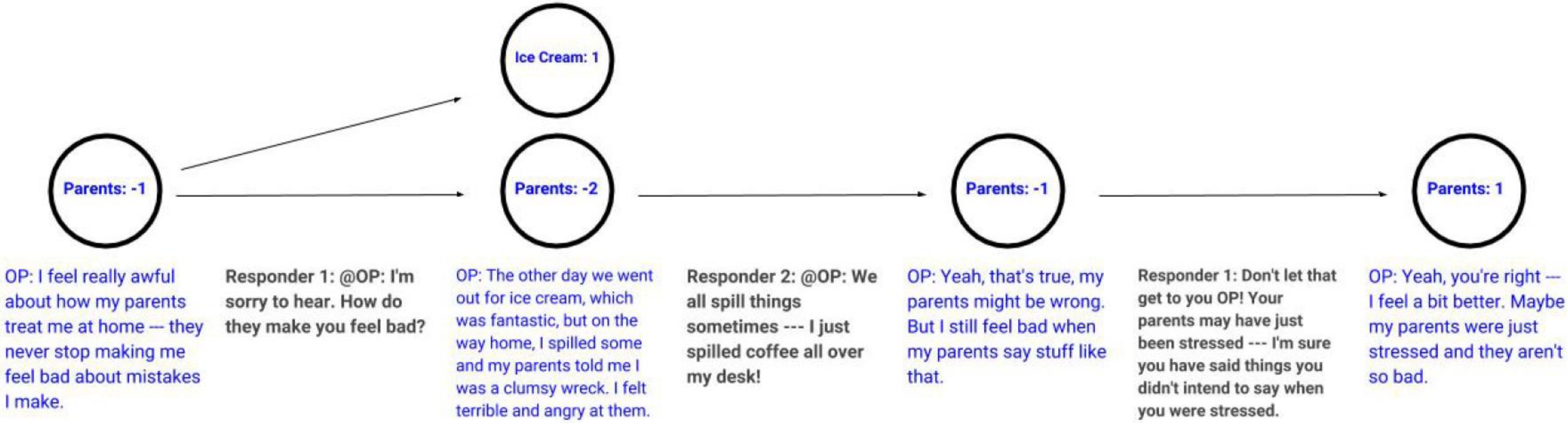
Results From Predictive Models (Cont'd)

Post-level AUC

	LIWC	LIWC + Punctuation	LIWC + Punctuation + Metadata	LIWC + Punctuation + Metadata + Language
CA Dataset	0.91	0.915	0.9	0.92
CA Dataset, only non-OP posts	0.92	0.926	0.923	0.92
CA Dataset, only OP posts	0.7	0.72	0.91	0.93
Indian Dataset	0.92	0.92	0.947	0.91
Indian Dataset, only non-OP posts	0.92	0.92	0.929	0.947
Indian Dataset, only OP posts	0.669	0.74	0.9645	0.927

Table 4: Post-level AUC for models trained on the Culturally Agnostic and the Indian dataset. We obtain higher AUC scores (> 0.9) than the thread-level models.

The SentiTopic Model



Algorithm 1 Extract-Topics Algorithm

1: For every noun n_j in each post, $\phi_j \leftarrow Sense2Vec(n_j), n_j \in Nouns$
2: Create clusters of k-nearest nouns for each distinct noun.
3: Repeat until convergence:
 •Merge similar clusters (avg. similarity $< \tau$)
 •Remove dissimilar words within each cluster (avg. similarity $> \tau$)