

Empathy Alignment in Online Communities

Honors Capstone
Jiamin Yang, advised by Prof. David Jurgens

Introduction

- More people are nowadays talking about their hard times on social media to seek for comfort and advice. Questions are left in how we should response to those posts that will be empathetic enough to be helpful.
- The Appraisal Theory is applied to characterize the trigger of emotions into several dimensions that help find the alignment between expressions, showing the sign of a person feeling the same as the other, namely empathy.

Background: The Appraisal Theory

- An appraisal is a description of one aspect of one's experience:
 1. Pleasantness: How pleasant the situation was.
 2. Anticipated Effort: How much effort was needed to deal with the situation.
 3. Situational Control: How much the situation was out of anyone's control.
 4. Self-Other Agency: How much oneself or another person was responsible for the situation.
 5. Attentional Activity: How much one's attention was drawn to rather than diverted away from the situation.
 6. Certainty: How certain about what was happening in the situation or what would happen next.
- Define target as the person experiencing an event or in a situation, and observer as the person who is responding to the target.
- Empathy occurs when the observer appraises the target's situation in the same way that the target appraises it.

Methods

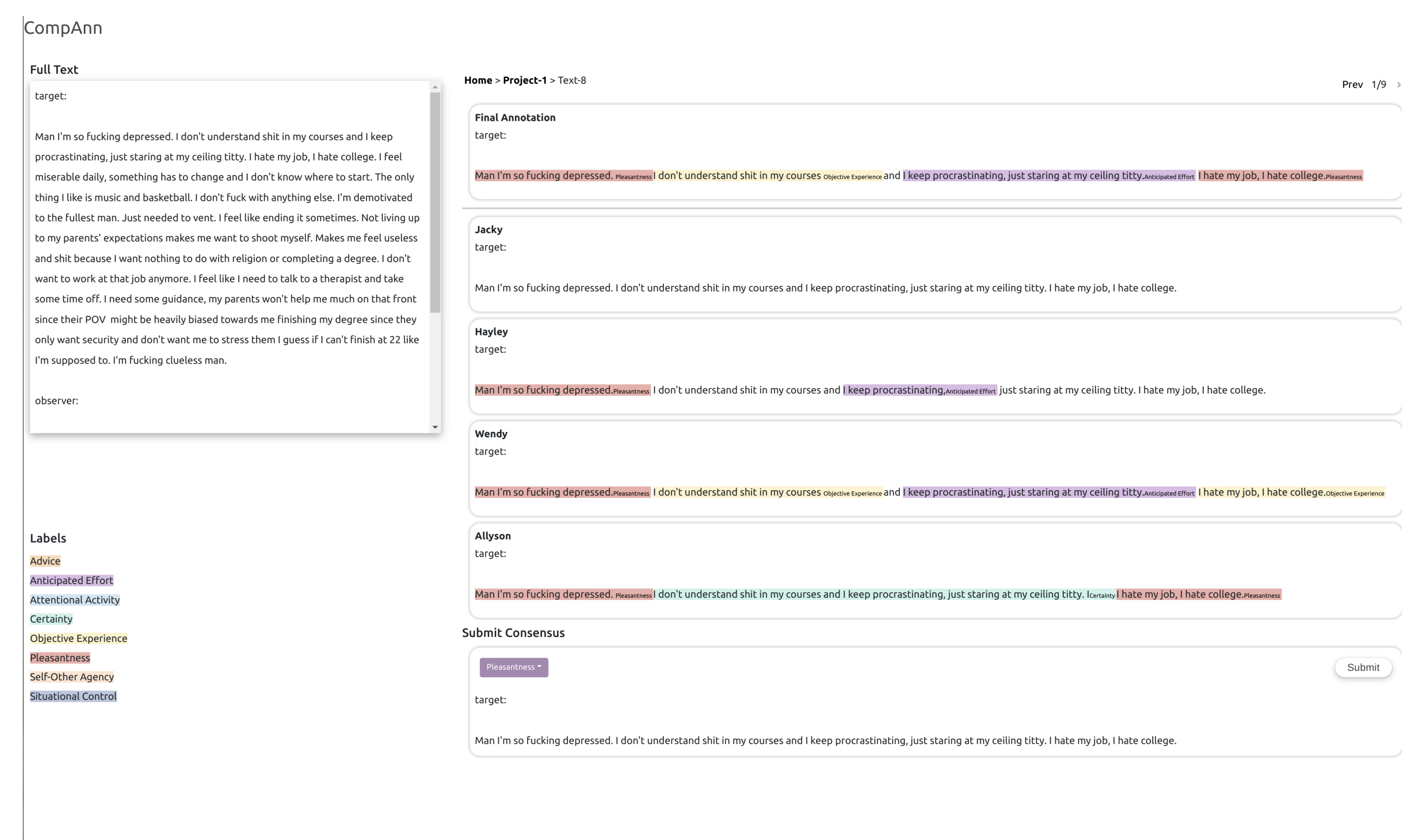
Data Collection

Paired Data with Target and Observer

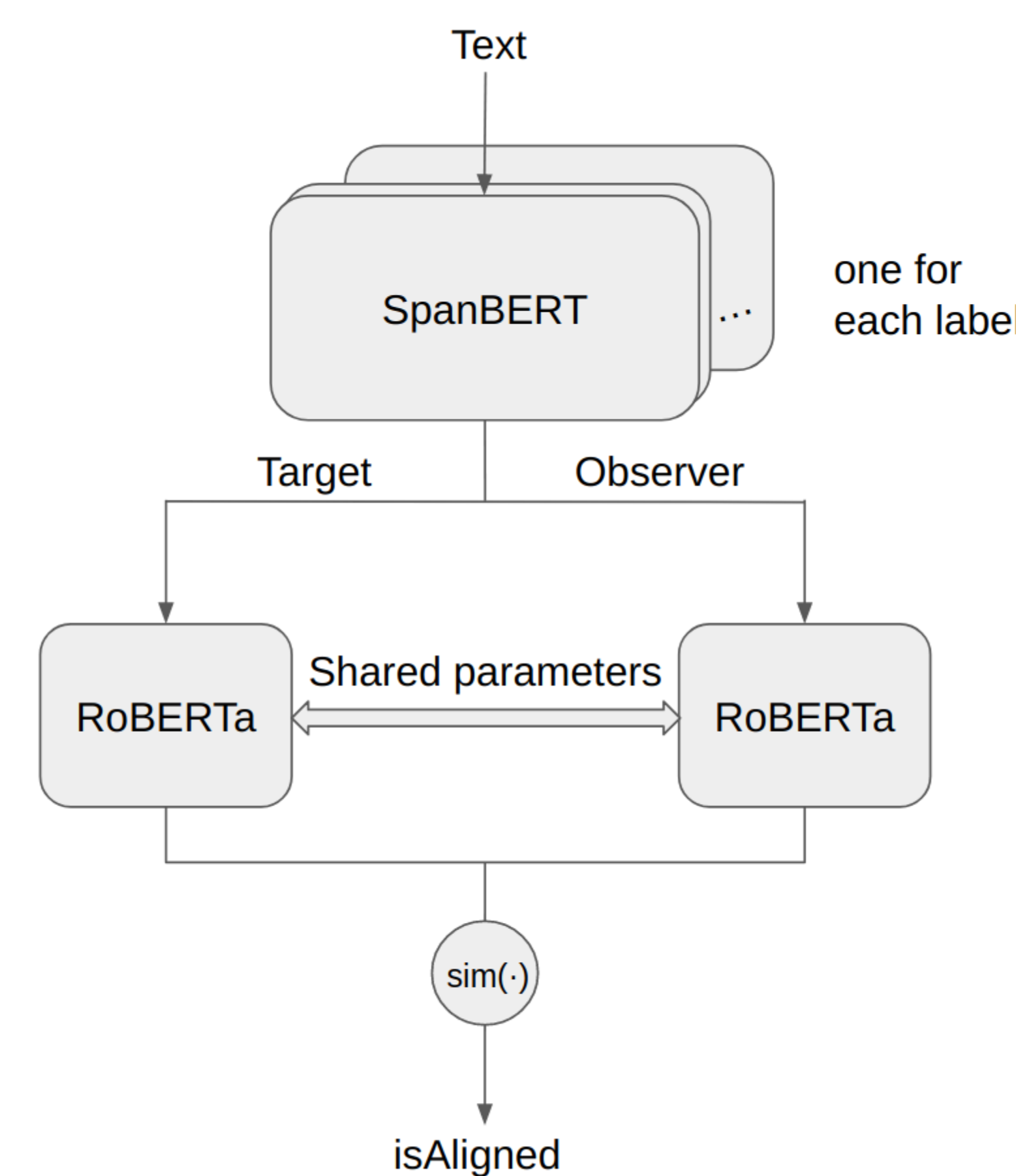
- Empathy Dataset with 246 pairs of target and observer text
- Reddit data from 2019.06 to 2021.06
 - 35 subreddits included (e.g. 'r/depression', 'r/heartbreak')
 - Two Classifiers are trained separately for filtering texts with high indication of distress or condolence using Microsoft/MiniLM-L12-H384-uncased.

Annotation

- Labels: Advice, Anticipated Effort, Attentional Activity, Certainty, Objective Experience, Pleasantness, Self-Other Agency, Situational Control
- Customized visualization tool for comparing annotations



Modeling



- SpanBERT: for recognizing each label from both target and observer text.
- Siamese Network with RoBERTa: for finding the alignment of feelings between the target and the observer in text spans that appraise in the same dimension.

Outcome

Classifier Performance

Classifier	Accuracy	F1
Distress	0.707	0.692
Condolence	0.799	0.79

Data Collected

- ~30M for one month of Reddit data
- 200+ annotated texts

Current Dataset Sample

- Pleasantness: "Gonna be real here...worst week of my life, haha."
- Anticipated effort: "I've been trying to be strong for so many people recently."
- Situational control: "This is my last chance to go to college and Uni"
- Self-other agency: "The little shit dropped me as a patient and now I can't seem to find a doctor who is willing to prescribe it."
- Attentional activity: "I am not typically an emotional person. I have only really and truly cried 3 times in my adult life, this being the 3rd and by far the worst."
- Certainty: "In retrospect, maybe should have broke it off earlier..."
- Objective Experience: "I just broke up with my girlfriend of 2 years like...a few days ago."
- Advice: "I highly recommend breathing and meditation type exercises."

Future Work

- Expand the current dataset and improve annotation codebook for clearer guidance.
- Try on heuristic methods for labeling dataset.
- Train and improve the designed model to check the performance and prepare for future analysis.