The Psychology of Semantic Spaces: Experiments with Positive Emotion (Student Abstract)

Xuan Liu[1] **Kokil Jaidka** [2] **Niyati Chhaya**[3]

[1] University of California Berkeley, [2] National University of Singapore, [3] Adobe Research
lxstephenlaw@gmail.com, jaidka@nus.edu.sg, nchhaya@adobe.com

Abstract

Psychological concepts can help computational linguists to better model the latent semantic spaces of emotions, and understand the underlying states motivating the sharing or suppressing of emotions. This abstract applies the understanding of agency and social interaction in the happiness semantic space to its role in positive emotion. BERT-based fine-tuning yields an expanded seed set to understand the vocabulary of the latent space. Results benchmarked against many emotion datasets suggest that the approach is valid, robust, offers an improvement over direct prediction, and is useful for downstream predictive tasks related to psychological states.

Introduction

When human beings post messages on social media, their emotive expressions provide a way to understand their well-being. This is because of the close association between moods and psychological states. Emotions such as joy, anger, fear, and sadness, are associated with innate psychological states such as happiness, well-being, depression, and stress. We turn this argument on its head and examine whether these psychological constructs can better describe the semantic spaces of emotion.

In considering the linguistic indicators of emotions, it is perhaps worthwhile to consider that they reflect how authors appraise themselves and their social environment. Psychologists define agency as the feeling of being in control of one's life. Its linguistic correlates have been examined in prior work (Jaidka et al. 2018). It is related to the ideas of autonomy (Tay and Diener 2011), and self efficacy (Deci and Ryan 2000), which are known to have a strong relationship with personal health and well-being (Lachman and Weaver 1998). On the other hand, social interaction and feeling connected to others are also central to well-being (Helliwell and Putnam 2004), a feeling of belonging (Sandstrom and Dunn 2014), and happiness (Epley and Schroeder 2014). These ideas are potentially useful to build a deeper semantic understanding of the language of emotion.

Herein lies the research gap, because the state of the art in emotion classifiers mainly rely on the use of valenced words or word contexts. A departure from this body of work is the CL-Aff HappyDB dataset (Jaidka et al. 2020), which explored and validated the use of agency and social interaction to understand happiness. Other work by Buechel et al. (2018) elaborates on "the psychological complexity" of human reactions such as empathy and distress, by annotating text data with the empathy assessments of their authors via multi-item scales. We can apply these ideas to better understand the emotion expressed in text as an offshoot of a psychological state. The benefit of applying psychological constructs to understanding evoked emotion would be a decreased reliance on affective words, and a finer sensitivity to the latent constructs in the semantic spaces.

We propose a generalized framework to map psychological constructs into the latent space of emotions. Inspired by work which uses task-specified knowledge to help the classification model improve the accuracy (Rozental, Kelrich, and Fleischer 2019), we first use weakly supervised learning methods to generate labels regarding the psychological constructs evident in text datasets with pre-existing emotion labels. Next, we concatenate this as the latent feature space into the encoder embedding of the original model trained on emotion labels. Then, we report preliminary experiments that validate our approach on four different emotion datasets. Finally, we demonstrate the performance improvements offered by such an approach to understand happiness.

Problem Formulation

Our latent variable framework aims at integrating specific psychological features into the original classification model to better clarify emotion. The challenge herein is that existing datasets do not have labels for psychological constructs; however, this is easily fixed by using weakly supervised methods. As the first step, we train models on the vocabulary of the CL-Aff HappyDB dataset to label other similar emotion datasets with agency and social interaction labels. We trained a multilayer perceptron to help us get the distribution of these psychological constructs. By splicing the psychological constructs distribution into the encoder embedding of the original model, we can then get a text representation oriented around the psychological constructs.

The basic approach to predict positive emotion involves using a BERT-based classifier which takes each observation and outputs the probability of a label. We have used a multimodal framework to incorporate categorical and numberical features in the model along with the email text. Besides the BERT embeddings, the psychological labels are also concatenated to the 768 dimension vector outputted by the BERT model and fed into a classifier using the Multimodal Toolkit.[1]

Results

Weakly Supervised Learning of Agency and Social Interaction

The labeled CL-Aff HappyDB dataset were used to train and validate classifiers on personal agency and social interaction in 10,000 autobiographical posts about happiness. Classifiers trained using multilayer perceptrons outperformed simpler approaches reported in (Jaidka et al. 2020) (Agency accuracy = 0.95, Social interaction accuracy = 0.88) and were subsequently applied to label an additional 87,622 posts that were similarly collected but did not include labels. They were also used to label four other emotion datasets:

- Sentiment Analysis (Emotion in Text) : It is a sentiment analysis dataset published on Kaggle. It includes 20,267 labeled observations collected through Crowdflower.
- International Survey on Emotion Antecedents and Reactions(ISEAR): It is the dataset that contains 7,672 observations collected by a large group of psychologists.
- Stance Sentiment Emotion Corpus (SSEC): The SSEC corpus is an annotation of the SemEval 2016 Twitter stance and sentiment corpus which contains 4,871 labeled observations.

External Validation

External validation suggested that incorporating agency and social interaction in predicting positive emotion has yielded improvements in the end-to-end BERT framework, as seen in Table 1.

Method	**Emotions in Text**	**ISEAR**	**SSEC**
MLPclassifier	80.66	92.23	68.79
MLPclassifier + A + S	**80.81**	**94.40**	68.38
Bert Finetune	84.31	97.40	77.41
Bert Finetune + A + S	**84.61**	**97.66**	77.41

Table 1: Predictive performance on positive emotion tasks. MLPclassifier refers to Multi-layer Perceptron classifier. A and S refer to Agency and Social interaction respectively.

Downstream Utility

On the CLAff HappyDB dataset, the results in Table 2 demonstrate that the final trained classifiers with agency and social interaction labels (Bert + Agency + Social interaction) were useful at predicting the 'duration' of happiness.[2]

[1] https://github.com/georgian-io/Multimodal-Toolkit

[2] 1 = "All day, I'm still feeling it" and 0 = "A few moments," "A few minutes," "At least one hour," and "Half a day"

Method	**Duration**
MLPclassifier	66.59
MLPclassifier + Agency + Social interaction	**67.38**
Bert Finetune	72.47
Bert + Agency + Social interaction	**72.52**

Table 2: Predictive performance for the duration of happiness in the CLAff HappyDB dataset, reported as Accuracy.

Future Work

In future work, we plan to generalize our framework to other emotions. Currently we are using low-dimensional representations of agency and social interaction and we plan to explore other ways to represent them. We plan to test the internal validity of our weakly supervised approach and offer more direct comparisons against self-reported psychological states.

References

Buechel, S.; Buffone, A.; Slaff, B.; Ungar, L.; and Sedoc, J. 2018. Modeling Empathy and Distress in Reaction to News Stories. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 4758–4765.

Deci, E. L.; and Ryan, R. M. 2000. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4): 227–268.

Epley, N.; and Schroeder, J. 2014. Mistakenly seeking solitude. *Journal of Experimental Psychology: General*, 143(5): 1980.

Helliwell, J. F.; and Putnam, R. D. 2004. The social context of well-being. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1449): 1435.

Jaidka, K.; Buffone, A.; Giorgi, S.; Eichstaedt, J.; Rouhizadeh, M.; and Ungar, L. 2018. Modeling and Visualizing Locus of Control with Facebook Language. In *Proceedings of the International AAAI Conference on Web and Social Media*.

Jaidka, K.; Chhaya, N.; Mumick, S.; Killingsworth, M.; Halevy, A.; and Ungar, L. 2020. Beyond Positive Emotion: Deconstructing Happy Moments Based on Writing Prompts. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 294–302.

Lachman, M. E.; and Weaver, S. L. 1998. The sense of control as a moderator of social class differences in health and well-being. *Journal of personality and social psychology*, 74(3): 763.

Rozental, A.; Kelrich, Z.; and Fleischer, D. 2019. Latent Universal Task-Specific BERT. arXiv:1905.06638.

Sandstrom, G. M.; and Dunn, E. W. 2014. Social interactions and well-being: The surprising power of weak ties. *Personality and Social Psychology Bulletin*, 40(7): 910–922.

Tay, L.; and Diener, E. 2011. Needs and subjective well-being around the world. *Journal of personality and social psychology*, 101(2): 354.