A Japanese Dataset for Subjective and Objective Sentiment Polarity Classification in Micro Blog Domain

Haruya Suzuki[†] Yuto Miyauchi[†] Kazuki Akiyama[†]
Tomoyuki Kajiwara[†] Takashi Ninomiya[†]
Noriko Takemura[‡] Yuta Nakashima[‡] Hajime Nagahara[‡]

[†] Graduate School of Science and Engineering, Ehime University, Japan
[‡] Institute for Datability Science, Osaka University, Japan

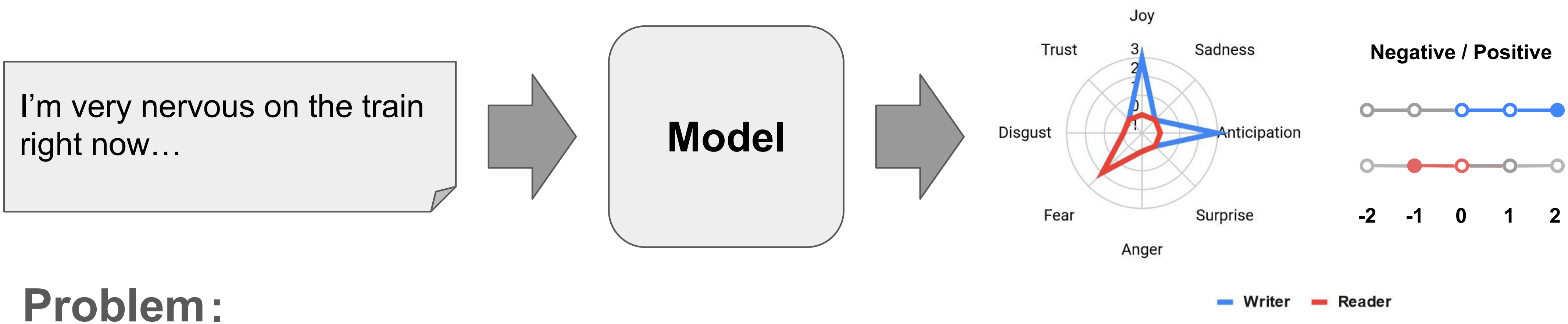
E-mail：suzuki@ai.cs.ehime-u.ac.jp

LREC 2022

Introduction (1/2)

Emotion Analysis:

- This task is to predict **emotion intensity** or **sentiment polarity** from text.



Problem:

- The relationship between emotion and sentiment or subjective and objective are not clear.

Introduction (2/2)

We extended the WRIME (Kajiwara+ 2021) dataset.

- Annotation : Sentiment polarity
- Number of annotators : 50 writers → 60 writers
- Dataset size : 17k posts → 35k posts

	Sentiment	Emotion	Subj.	Obj.	Langage	Size
IMDB (Maas+ 2011)	✓	×	✓	×	English	50,000
SST (Socher+ 2013)	✓	×	×	✓	English	11,855
SemEval-2018 (Mohammad+ 2011)	×	✓	×	✓	English	12,634
WRIME (Kajiwara+ 2021)	×	✓	✓	✓	Japanese	17,000
Ours	✓	✓	✓	✓	Japanese	35,000

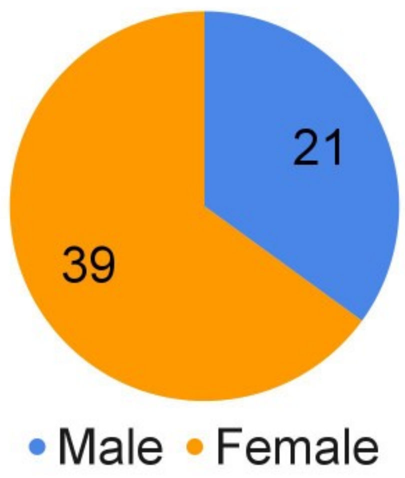
Table of Contents

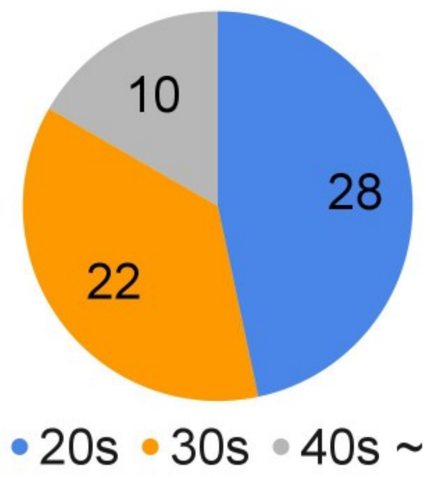
Annotating Subjective Labels

We hired 60 participants via crowdsourcing service.

- *Lancers*: https://www.lancers.jp/
- They annotated subjective labels their own SNS posts.
  - Sentiment Polarity
  - Five-point scale (-2・-1・0・1・2)

Negative ← Neutral → Positive





Annotating Objective Labels

We hired 3 annotators via crowdsourcing service.

- *Lancers*: https://www.lancers.jp/
- They annotated objective labels for all posts.
  - Objective labels are estimated writer's emotions.
  - Two woman in their 30s and one woman in their 40s.

I'm very nervous on the train right now…

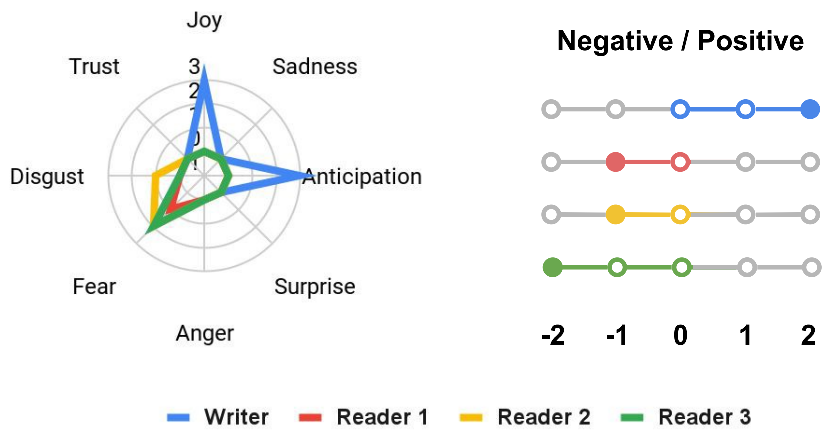


Table of Contents

Distribution of Sentiment Polarity

- Reader 1 labeled more positives and negatives than neutrals.
- Reader 2 labeled more neutral labels and less extreme labels.
- Reader 3 labeled more strong positives and strong negatives.
- Avg. Readers fewer extreme strong positive and strong negative and more negative labels.

	-2	-1	0	1	2
Writer	4,105	6,465	10,380	9,415	4,635
Avg. Readers	1,687	10,468	11,462	9,138	2,245
Reader 1	2,254	10,316	8,741	11,216	2,473
Reader 2	1,056	4,029	20,147	8,510	1,258
Reader 3	9,581	4,256	10,687	2,841	7,635

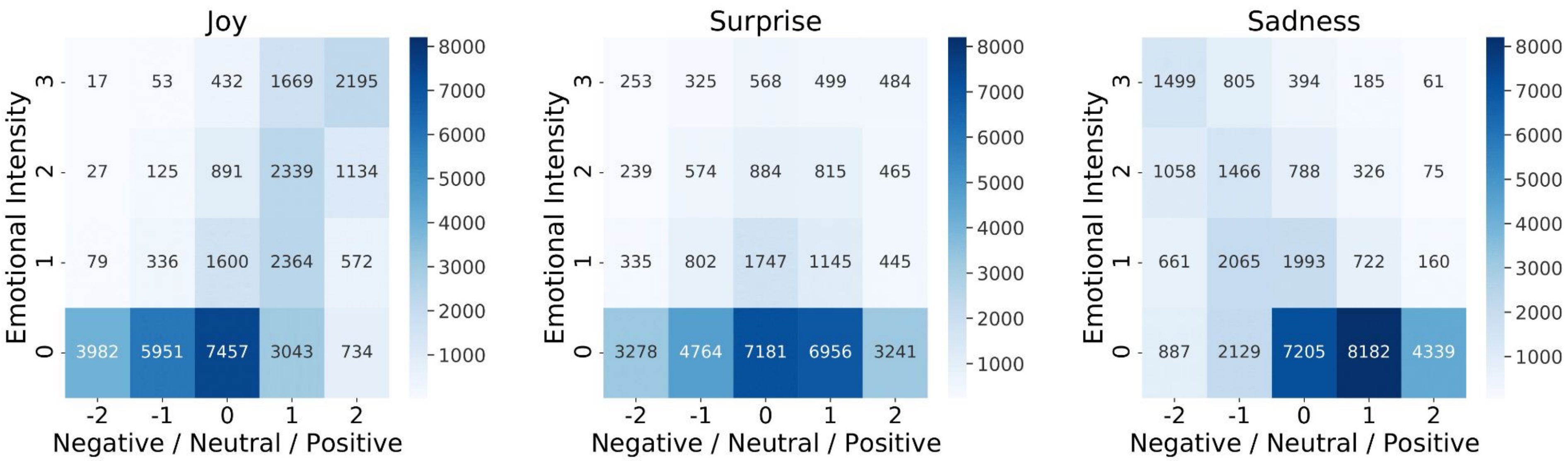
Pearson Correlation between Emotion & Sentiment

- *Joy*, *Anticipation*, and the *Trust*, are positive emotions.
- *Sadness*, *Anger*, *Fear*, and the *Disgust*, are negative emotions.
- *Surprise* is neutral emotion.

	Joy	Anticipation	Trust	Surprise	Sadness	Anger	Fear	Disgust
Writer	0.585	0.381	0.296	0.052	-0.526	-0.353	-0.298	-0.467
Avg. Readers	0.665	0.400	0.252	0.037	-0.539	-0.229	-0.410	-0.470

Distribution of Emotion & Sentiment

- *Joy* is a positive correlation between emotion and sentiment.
- *Sadness* is a negative correlation between emotion and sentiment.
- *Surprise* appears in both positive and negative posts.



Agreement between Subjective & Objective

The agreement between sentiment polarity labels by subjective annotators and those by objective annotators.

- No emotion：All emotional intensities are none or weak.
- Single emotion：Only one emotional intensity is medium or strong.
- Multiple emotions：Two or more emotional intensities are medium or strong.

The posts show that includes multiple emotions indicate a higher agreement.

	#Posts	QWK
No emotion	11,395	0.496
Single emotion	12,281	0.590
Multiple emotions	11,324	0.697

Table of Contents

Sentiment Polarity Classification (1/2)

Models :

- Bag-of-Words (BoW) + Logistic regression (LogReg)
- BERT (Wikipedia) : A model pre-trained using Wikipedia.
- BERT (SNS) : A model pre-trained using SNS texts.
- Subj. BERT (SNS) : A model fine-tuned by the writer’s labels.
- Obj. BERT (SNS) : A model fine-tuned by the readers’s labels.

Metrics :

- Accuracy
- Mean Absolute Error (MAE)
- Quadratic Weighted Kappa (QWK)

	Train	Dev	Test
Writers	40	10	10
Posts	30,000	2,500	2,500

Sentiment Polarity Classification (2/2)

- The performance of Obj. BERT (SNS) is consistently high.
- Estimating the reader’s labels is easier than that of writers.

	Subjective			Objective		
	Accuracy	MAE	QWK	Accuracy	MAE	QWK
BoW+LogReg	0.344	0.924	0.359	0.443	0.695	0.444
BERT (Wikipedia)	0.386	0.824	0.512	0.573	0.483	0.695
BET (SNS)	0.391	0.778	0.558	0.615	0.426	0.743
Subj. BERT (SNS)	0.391	0.778	0.558	0.443	0.646	0.627
Obj. BERT (SNS)	0.436	0.694	0.595	0.615	0.426	0.743

Conclusion

We extended the WRIME (Kajiwara+ 2021) dataset.

https://github.com/ids-cv/wrime

- We annotated all of the **emotion intensity** and **sentiment polarity** or **subjective** and **objective**.
- We found that emotion-sentiment correlations and the text with multiple emotions were more likely to perceive the sentiment polarity of the writer.
- Experimental results on sentiment polarity classification show that it is more difficult to estimate the writer's subjective sentiment than the reader's objective ones.