Appendices

A Model Hyperparameters

We provide all the hyperparameters used to train our models in Table 1.

Type	Values	Models
# of Enriched Patterns	19,820	CNN_{BASIC}
	160,905	$CARER_{ch}$
	187,648	$CARER_{\beta}$
	476,174	CARER, $CARER_{EK}$
Size of Embedding	(10,000, 300)	RNN_{w2v}, Bi-GRNN
	(10,696, 300)	$fastText_{ch}$
	(80,000, 300)	EmoNet
	(1,859,184, 200)	CNN_{w2v}
	(3,000,000, 300)	CNN_{char}
Window Size	(3)	CNN_{w2v}
	(3, 16)	CARER, $CARER_{\beta}$, $CARER_{EK}$, CNN_{BASIC}
	(1, 3, 8)	$CARER_{ch}$, $fastText_{ch}$
	(3, 4, 5)	CNN_{char}
# of Filters	100	CNN_{char}
	256	CARER, $CARER_{\beta}$, $CARER_{EK}$, $CARER_{ch}$, CNN_{BASIC}, CNN_{w2v}, $fastText_{ch}$
Learning Rate	0.001	EmoNet, RNN_{w2v}, Bi-GRNN
Hid. Layer Dim.	(256)	RNN_{w2v}, Bi-GRNN
	(1024)	CNN_{w2v}
	(512, 128)	CARER, $CARER_{\beta}$, $CARER_{EK}$, $CARER_{ch}$, CNN_{BASIC}, $fastText_{ch}$
	(1000, 1000, 1000)	EmoNet
Conv Dropout	0.3	CARER, $CARER_{\beta}$, $CARER_{EK}$, $CARER_{ch}$, CNN_{BASIC}, CNN_{w2v}, $fastText_{ch}$
	0.5	EmoNet
Dense Dropout	0.2	RNN_{w2v}, Bi-GRNN
	0.5	CARER, $CARER_{\beta}$ $CARER_{EK}$, $CARER_{ch}$, CNN_{BASIC}, $fastText_{ch}$, EmoNet
Batch Size	50	CNN_{char}
	128	CARER, $CARER_{\beta}$, $CARER_{EK}$, $CARER_{ch}$, CNN_{BASIC}, $fastText_{ch}$, EmoNet, CNN_{w2v}
	256	RNN_{w2v}, Bi-GRNN
Epoch	3	$CARER_{\beta}$
	4	CARER
	7	EmoNet
	8	$CARER_{EK}$
	12	$CARER_{ch}$, $fastText_{ch}$, Bi-GRNN
	20	CNN_{BASIC}
	22	CNN_{w2v}
	24	RNN_{w2v}
	50	CNN_{char}

Table 1: Hyperparameters of proposed models.

B CARER for Traditional Chinese

CARER$_{ch}$ is a model trained to support emotion recognition on a Traditional Chinese dataset. The dataset is crawled from comments found in *Facebook Fanpages*. We selected 11,268 fan pages, which contain high engagement between Chinese-speaking Facebook users, on topics such as religion, politics, news press, YouTubers, parenting, life, comics, gaming, etc. Users' comments on fan pages are considered for the dataset, with emotion labels automatically determined by the *emoji reaction* of the commenting users on the root post. Only four emotions are considered: *anger*, *joy*, *sadness*, and *surprise*. A total of 10 million comments were retrieved and tokenized through point-wise mutual information. Comments containing less than four or greater than 100 tokens are purged, which resulted in a dataset of 718,852 comments. The dataset is split into a training and testing dataset using a 90/10 split. The emotion distribution of each dataset is provided below:

Training Dataset	
Emotion	**Amount**
joy	168,404
anger	168,203
surprise	159,989
sadness	150,370

Testing Dataset	
Emotion	**Amount**
joy	18,734
anger	18,601
surprise	17,791
sadness	16,760

Table 2: Data statistics of datasets used in **CARER$_{ch}$**.

Chinese-based enriched patterns are then derived using the proposed graph-based feature extraction mechanism. Word clusters are formed and validated through moedict-data [1]. These patterns are then used to train a CNN-based emotion recognition model, similar to the one used to conduct the English experiments. We observed that since tokens containing one to three characters, a flexible range of window sizes does help the classifier to capture important features used to distinguish statements conveying different emotions. Consequently, window sizes of $(1, 3, 8)$ are used for the Chinese-based emotion recognition experiment. Overall, we show that the proposed approach is not restricted to any specific language and that the enriched features are applicable to other languages and data sources.

[1] moedict-data is an open-sourced synonym dictionary available at https://github.com/g0v/moedict-data .

C English-based Emotion Lexicon

Note that a preview of the pattern-based emotion lexicon (English version) is provided as supplementary material. The zip folder (8emos.zip) contains 8 files (one for each emotion), each containing patterns and the corresponding pattern scores. (Licensed under The MIT License [2])

[2] https://opensource.org/licenses/MIT