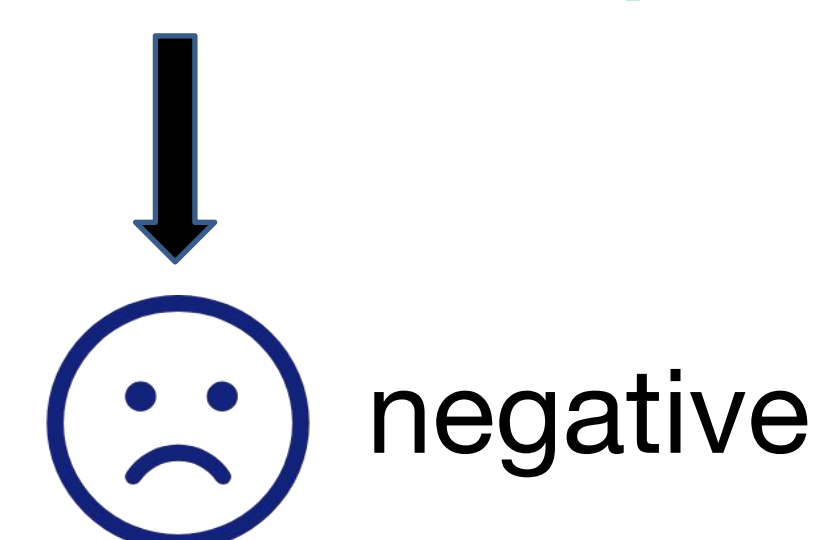


Targeted Sentiment Classification with Attentional Encoder Network

Youwei Song, Jiahai Wang*, Tao Jiang, Zhiyue Liu, and Yanghui Rao
Sun Yat-sen University, Guangzhou, China School of Data and Computer Science
Contact Author: Jiahai Wang (wangjiah@mail.sysu.edu.cn)

Targeted sentiment classification aims at determining the sentimental tendency towards specific targets. Most of the previous approaches model context and target words with RNN and attention. However, RNNs are difficult to parallelize and truncated backpropagation through time brings difficulty in remembering long-term patterns. To address this issue, this paper proposes an Attentional Encoder Network (AEN) which eschews recurrence and employs attention based encoders for the modeling between context and target. We raise the label unreliability issue and introduce label smoothing regularization. We also apply pre-trained BERT to this task and obtain new state-of-the-art results. Experiments and analysis demonstrate the effectiveness and lightweight of our model.

The **price** is **reasonable** although the **service** is **poor**.



Intra-MHA for introspective context words modeling.

Inter-MHA for context-perceptive target words modeling.

PCT: transform contextual information gathered by the MHA. Point-wise means that the kernel sizes are 1 and the same transformation is applied to every single token belonging to the input.

Models	Twitter		Restaurant		Laptop	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
TD-LSTM	0.7080	0.6900	0.7563	-	0.6813	-
ATAE-LSTM	-	-	0.7720	-	0.6870	-
IAN	-	-	0.7860	-	0.7210	-
RAM	0.6936	0.6730	0.8023	0.7080	0.7449	0.7135
Feature-based SVM	0.6340	0.6330	0.8016	-	0.7049	-
Rec-NN	0.6630	0.6590	-	-	-	-
MemNet	0.6850	0.6691	0.7816	0.6583	0.7033	0.6409
AEN-GloVe w/o PCT	0.7066	0.6907	0.8017	0.7050	0.7272	0.6750
AEN-GloVe w/o MHA	0.7124	0.6953	0.7919	0.7028	0.7178	0.6650
AEN-GloVe w/o LSR	0.7080	0.6920	0.8000	0.7108	0.7288	0.6869
AEN-GloVe-BiLSTM	0.7210	0.7042	0.7973	0.7037	0.7312	0.6980
AEN-GloVe	0.7283	0.6981	0.8098	0.7214	0.7351	0.6904
BERT-SPC	0.7355	0.7214	0.8446	0.7698	0.7899	0.7503
AEN-BERT	0.7471	0.7313	0.8312	0.7376	0.7993	0.7631

Models	Model size	
	Params $\times 10^6$	Memory (MB)
TD-LSTM	1.44	12.41
ATAE-LSTM	2.53	16.61
IAN	2.16	15.30
RAM	6.13	31.18
MemNet	0.36	7.82
AEN-BERT	112.93	451.84
AEN-GloVe-BiLSTM	3.97	22.52
AEN-GloVe	1.16	11.04

