LREC-COLING2024

Let's Rectify Step by Step: Improving Aspect-based Sentiment Analysis with Diffusion Models

Shunyu Liu, Jie Zhou, Qunxi Zhu, Qin Chen, Qingchun Bai, Jun Xiao, Liang He

East China Normal University
Computer Science and Technology

CONTENT

REPORT

01. Introduction

REPORT

Aspect term Extraction (AE) and Sentiment Classification (SC) :

Example: Amazing Spanish Mackeral special appetizer and perfect box sushi (that eel with avodcao – um um um).

Positive

Positive

Challenge:

Longer Terms(contain multiple words), ambiguity and fluidity of language usage, …

01. Introduction

REPORT

Diffusion Model:

1. Impressive capabilities in controlled generation tasks

2. Remarkable performance in various domains, including text-to-image , and text generation.

01. Introduction

REPORT

DiffusionABSA:

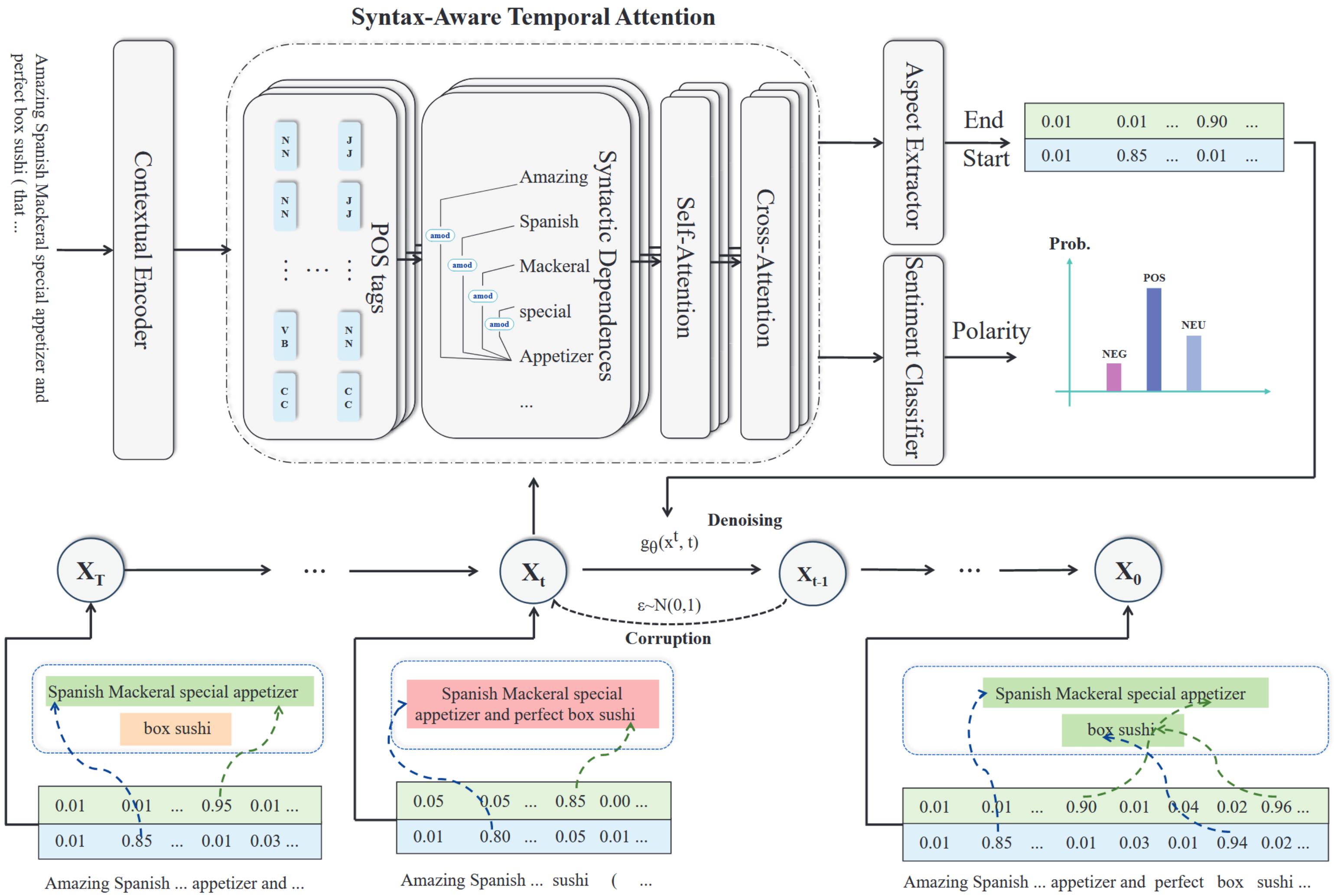
STEP 1	Amazing $\overbrace{\text{Spanish Mackeral special appetizer}}^{\text{ASPECT ✓ POSITIVE ✓}}$ and perfect $\overbrace{\text{box sushi}}^{\text{MISSING ✗}}$ (that eel with avodcao – um um um).
STEP 2	Amazing $\overbrace{\text{Spanish Mackeral special appetizer and perfect box sushi}}^{\text{ASPECT ✗ POSITIVE ✓}}$ (that eel with avodcao – um um um).
STEP 3	Amazing $\overbrace{\text{Spanish Mackeral special appetizer}}^{\text{ASPECT ✓ POSITIVE ✓}}$ and perfect $\overbrace{\text{box sushi}}^{\text{ASPECT ✓ POSITIVE ✓}}$ (that eel with avodcao – um um um).
	...

Table 1: The boundary of aspect terms gradually changes during the denoising process in DiffusionABSA. The spans annotated with green, orange, and red respectively signify the correct, missing, wrong results.

02. Methods

REPORT

Framework of DiffusionABSA:



03. Experiments

REPORT

Datasets:

Restaurant and laptop reviews in two version

Dataset	14res		14lap		15res		16res	
	#S	#T	#S	#T	#S	#T	#S	#T
				D_{20a}				
Train	1300	2145	920	1265	593	923	842	1289
Dev	323	524	228	337	148	238	210	316
Test	496	862	339	490	318	455	320	465
				D_{20b}				
Train	1266	2338	906	1460	605	1013	857	1394
Dev	310	577	219	346	148	249	210	339
Test	492	994	328	543	322	485	326	514

Table 2: Statistics of the datasets pertinent to the ABSA task. #S and #T denote the quantities of sentences and targets within the respective datasets.

03. Experiments

REPORT

Effectiveness of Our Model:

MODEL	14res	14lap	15res	16res
CMLA+	70.62	56.90	53.60	61.20
RINANTE+	48.15	36.70	41.30	42.10
Li-unified	73.79	63.38	64.95	70.20
Peng-two-stage	74.19	62.34	65.79	71.73
Dual-MRC	76.57	64.59	65.14	70.84
SPAN-BART	78.47	68.17	69.95	75.69
SyMux	78.68	70.32	69.08	77.95
SynGen	79.72	70.06	71.61	77.51
ChatGPT (Zero-shot)	59.08	45.48	53.91	55.40
ChatGPT (5-shot ICL)	65.98	49.50	63.66	63.11
ChatGPT (5-shot COT)	62.82	48.87	66.07	65.93
DiffusionABSA	**80.93**	**72.81**	**76.70**	**81.72**
w/o SynTA	80.84	72.39	74.26	80.53

Table 3: Results of AESC over D_{20a} datasets. We use the results of baselines reported in Yu et al. (2023).

MODEL	14res	14lap	15res	16res
SPAN	86.71	82.34	74.63	74.68
RACL	86.38	81.79	73.99	74.91
MIN	87.91	83.22	-	-
CMLA	81.22	79.53	76.03	74.20
RINANTE	81.34	80.40	73.38	72.82
Li-unified	81.62	78.56	74.65	73.36
GTS	83.82	82.48	78.22	75.80
GEN	87.07	83.52	75.48	81.35
SyMux	**89.02**	84.42	79.73	82.41
ChatGPT (Zero-shot)	55.65	43.03	40.33	-
ChatGPT (5-shot ICL)	70.99	48.19	53.49	-
ChatGPT (5-shot COT)	72.41	54.50	59.27	-
DiffusionABSA	87.15	**86.66**	**85.40**	**87.87**
w/o SynTA	86.93	86.01	83.15	86.23

Table 4: Results of AE over D_{20a} datasets. We use the results of baselines reported in Fei et al. (2022).

MODEL	14res		14lap		15res	
	AE	AESC	AE	AESC	AE	AESC
SPAN-BERT	86.71	73.68	82.34	61.25	74.63	62.29
IMN-BERT	84.06	70.72	77.55	61.73	69.90	60.22
RACL-BERT	86.38	75.42	81.79	63.40	73.99	66.05
Dual-MRC	86.60	75.95	82.51	65.94	75.08	65.08
DiffusionABSA	**86.17**	**80.64**	**88.37**	**74.90**	**84.62**	**77.26**
w/o SynTA	85.89	80.11	84.74	70.40	85.62	77.02

Table 6: Results of AE, AESC on D_{20b} datasets.

03. Experiments

REPORT

Effectiveness of Our Model:

MODEL	14res				16res			
	ALL	LEN=1	LEN=2	LEN>2	ALL	LEN=1	LEN=2	LEN>2
SeqLab	66.17	58.88	16.11	4.36	68.60	58.94	19.37	4.73
DiffusionABSA	79.13	71.68	20.82	7.51	78.87	68.97	21.61	8.49
Improvement	19.59%	21.74%	29.24%	72.25%	14.97%	17.02%	11.56%	79.49%

Table 5: Results over aspects with various lengths on D_{20a} datasets.

04. Conclusions

REPORT

Conclusions:

- We propose DiffusionABSA, a novel framework that adapts diffusion models to refine the aspect progressively through a dynamic interplay of corruption and denoising processes.
- We design a denoising neural network enhanced by a syntax-aware temporal attention module.
- A series of experiments on eight widely-used benchmark datasets show that DiffusionABSA achieves new SOTA performance in most cases.

Thanks!

Q&A

Feel free to contact me (71265901045@stu.ecnu.edu.cn)