LREC-COLING 2024

FUDAN UNIVERSITY

Linking Adaptive Structure Induction and Neuron Filtering: A Spectral Perspective for Aspect-based Sentiment Analysis

Hao Niu*†, Maoyi Wang*, Yun Xiong, Biao Yang, Xing Jia, Zhonglei Guo
Shanghai Key Laboratory of Data Science, School of Computer Science, Fudan University
Shanghai, China
{hniu18, wangmaoyi, yunx, xjia18, guozl18}@fudan.edu.cn
biaoyang22@m.fudan.edu.cn

Background

- Sentence (S): The **decor** is not a special at all but their amazing **food** makes up for it.
- Aspects (A):
  - **decor** -> NEGATIVE
  - **food** -> POSITIVE

- **ABSA: Aspect-based sentiment analysis**

Related Works

- Most approaches based on GNNs (e.g. ASGCN, RGAT, etc.);

- **i) External structure,**
- **ii) Semi-induced structure,**
- **iii) Full-induced structure.**

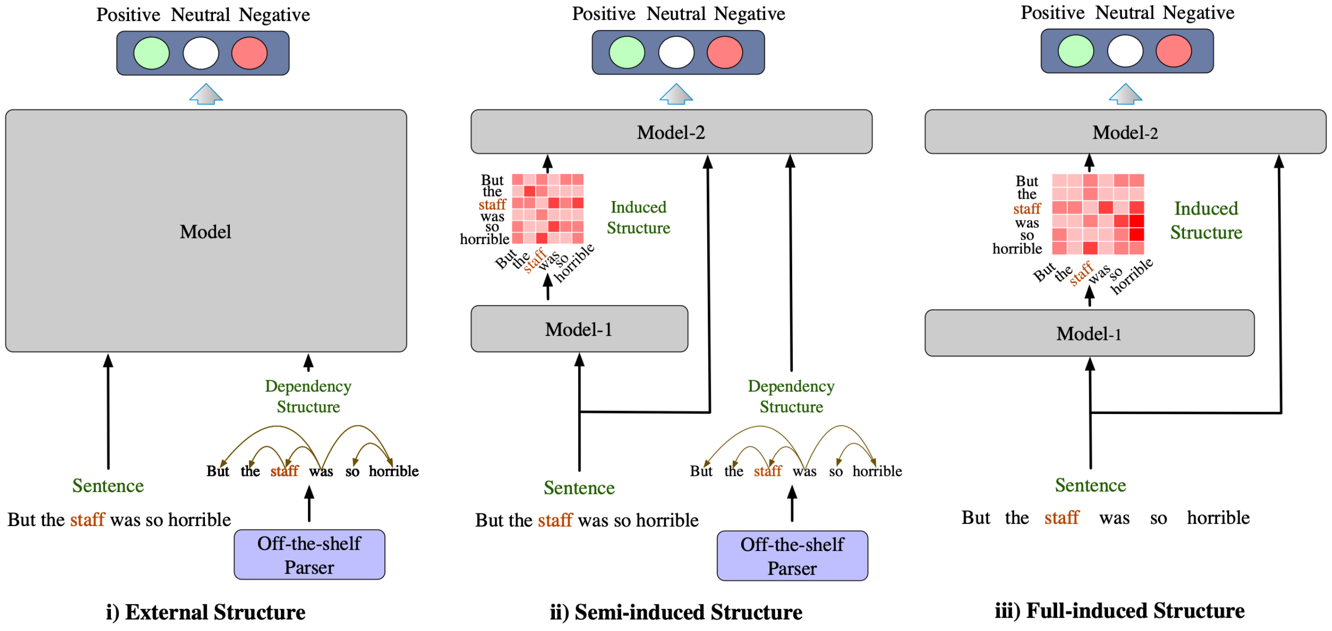

Figure 1: Taxonomy of structure-based ABSA.

Motivations

- Following the paradigm iii) Full-induced structure, to perform Structure Induction for ABSA
- Our research aims to explore the impact of neuron-level manipulation on structure induction for ABSA

Contributions

- **Neuron-level Manipulations**. Neuron-level manipulations can influence structure induction. The induced structures of NeuLT obtain lower AsD and better performance compared to the Attn. method
- **Structure Induction.** GSL-based structure induction is effective. The Attn. is more suitable for structure induction compared to Kernel and Cosine.
- **Extensive Experiments and Neuron-level Analysis.** We conduct extensive experiments and analysis. Results confirm our findings and demonstrate the effectiveness of NeuLT, and neuron-level analysis provides in-depth insights into the approach.

Approach

- **Graph Structure Learning (GSL)**

$$\epsilon_{ij} = (\mathbf{W}_i \mathbf{h}_i)(\mathbf{W}_j \mathbf{h}_j)^\top,$$

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ \frac{exp(\epsilon_{ij})}{\sum_{k=1}^{n} exp(\epsilon_{ik})} & \text{otherwise} \end{cases},$$

$$\mathbf{H}^l = \sigma\left(\mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \mathbf{H}^{l-1} \mathbf{W}^l\right),$$

- **Neuron Filtering (NeuLT)**

$$\Upsilon^{nlt}(x) = \mathcal{F}^{-1}\left(\Pi(\mathcal{F}(x))\right),$$

$$\epsilon_{ij} = \Upsilon^{nlt}(\mathbf{W}_i \mathbf{h}_i)\Upsilon^{nlt}(\mathbf{W}_j \mathbf{h}_j)^\top,$$

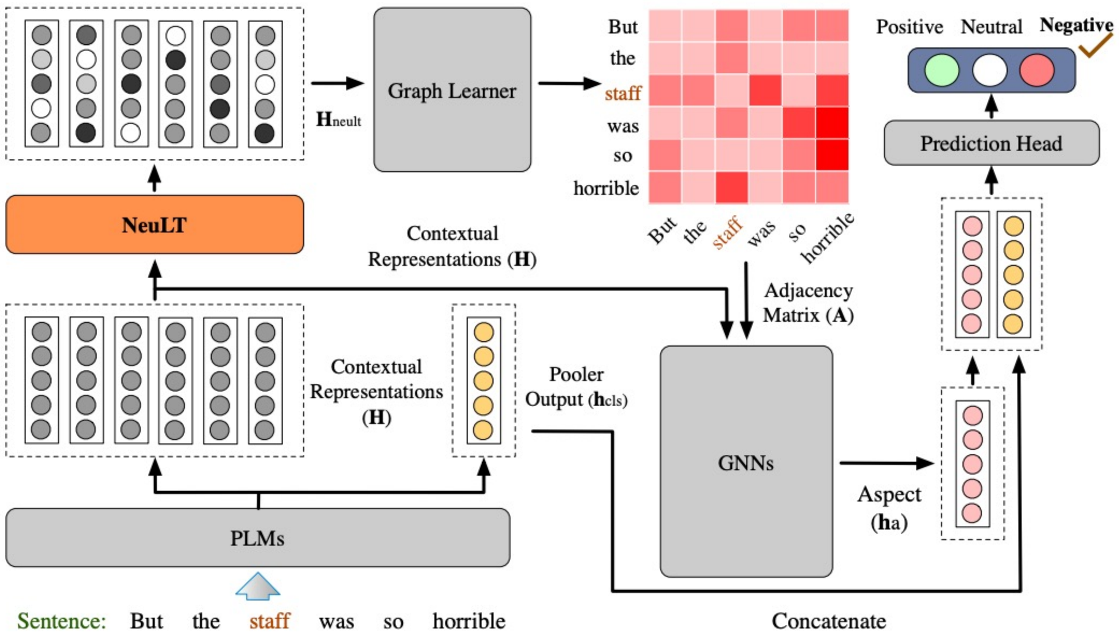

Figure 2: Overall architecture. The aspect (staff) with NEGATIVE sentiment polarity label is in red.

Datasets

- Statistics of three datasets

Table 1: Statistics of datasets.

Dataset	Positive		Neutral		Negative	
	Train	Test	Train	Test	Train	Test
Rest14	2164	728	807	196	637	196
Laptop14	994	341	870	128	464	169
Twitter	1561	173	3127	346	1560	173

Overall Performance

Table 2: The overall performance across the three datasets. The baselines in the 'Structure' column are classified according to the structure categorization (*Dep.*: external structures (dependency syntactic tree), *Semi.*: semi-induced structures, *Full*: full-induced structures, and *None*: no structure information used).

Embedding	Model	Structure	Rest14		Laptop14		Twitter	
			Accuracy	*Macro-F1*	*Accuracy*	*Macro-F1*	*Accuracy*	*Macro-F1*
Static Embedding	depGCN	Dep.	80.77$^\sharp$	72.02$^\sharp$	75.55$^\sharp$	71.05$^\sharp$		
	CDT	Dep.	82.30$^\sharp$	74.02$^\sharp$	77.19$^\sharp$	72.99$^\sharp$		
	kumaGCN	Semi.	81.43	73.64	76.12	72.42	72.45	70.77
	RGAT	Dep.	83.30	76.08	77.42	73.76	75.57	73.82
	FT-RoBERTa(ASGCN)	Full	82.31	73.53	76.33	72.76	73.84	72.66
	FT-RoBERTa(PWCN)	Full	82.40	73.95	76.95	73.21	73.84	71.43
	FT-RoBERTa(RGAT)	Full	82.76	75.25	77.43	74.21	75.43	74.04
BERT$_{base}$	BERT	None	85.62$^\sharp$	78.28$^\sharp$	77.58$^\sharp$	72.38$^\sharp$	75.28	74.11
	SAGAT	Dep.	85.08	77.94	80.37	76.94	75.40	74.17
	DGEDT	Dep.	86.30	80.00	79.80	75.60	77.90	75.40
	depGCN-BERT	Dep.	85.00	78.79	81.19	77.67	75.58	74.58
	RGAT-BERT	Dep.	86.60	81.35	78.21	74.07	76.15	74.88
	KumaGCN-BERT	Semi.	86.43	80.30	81.98	78.81	77.89	77.03
	dotGCN-BERT	Full	86.16	80.49	81.03	78.10	78.11	77.00
RoBERTa$_{base}$	Roberta + MLP	None	87.32	81.01	82.60	79.33	77.17	76.20
	RoBERTa-ASC(Dep)	Dep.	82.82	75.12	74.12	70.52	-	-
	LCFS-ASC-CDW(Dep)	Dep.	86.71	80.31	80.52	77.13	-	-
	Dep(ASGCN)	Dep.	86.90	80.75	81.66	78.31	75.28	74.38
	Dep(PWCN)	Dep.	87.41	81.07	84.16	81.18	76.63	75.60
	Dep(RGAT)	Dep.	87.43	80.61	83.43	80.28	74.42	72.93
	FT-RoBERTa(ASGCN)	Full	86.87	80.59	83.33	80.32	76.10	75.07
	FT-RoBERTa(PWCN)	Full	87.35	80.85	84.01	81.08	77.02	75.52
	FT-RoBERTa(RGAT)	Full	87.52	81.29	83.33	79.95	75.81	74.91
	NeuLT	Full	88.93	83.28	84.95	82.26	78.18	77.59
RoBERTa$_{large}$	**NeuLT**	Full	**89.64**	**84.18**	**86.05**	**84.68**	**78.53**	**77.78**

LREC-COLING 2024

Ablation Study

Table 3: Results of Ablation Studies.

Embedding	Model	Structure	Rest14		Laptop14		Twitter	
			Accuracy	*Macro-F1*	*Accuracy*	*Macro-F1*	*Accuracy*	*Macro-F1*
BERT_{base}	*Attn.*	Full	85.43	78.04	80.54	77.06	76.22	75.04
	NeuLT	Full	**86.95**	**81.20**	**81.33**	**77.20**	**77.10**	**75.83**
RoBERTa_{base}	*Attn.*	Full	87.59	81.72	83.86	80.53	75.72	73.92
	NeuLT	Full	**88.93**	**83.28**	**84.95**	**82.26**	**78.18**	**77.59**
RoBERTa_{large}	*Attn.*	Full	89.46	84.12	84.80	82.19	77.02	75.75
	NeuLT	Full	**89.64**	**84.18**	**86.05**	**84.68**	**78.53**	**77.78**

Aspects-sentiment Distance (AsD)

Table 5: The Aspects-sentiment Distance (AsD) across different structures in all datasets. The dependency tree structure (Dep.) is derived from the Spacy parser.

Structure	Rest14	Laptop14	Twitter
Dep.	8.19	8.02	8.33
Attn.	2.26	2.55	2.64
NeuLT	**2.04**	**2.39**	**2.48**

Automatic Neuron Filtering (NeuLT(Auto))

- **Method Description.**

$$\Pi^{ans}(\mathcal{F}(\mathbf{H})) = \mathbf{M}_{bern} \odot \mathbf{H}_{fft},$$

$$\Upsilon^{afs}(x) = \mathcal{F}^{-1}\Big(\Pi^{ans}\big(\mathcal{F}(x)\big)\Big),$$

$$\epsilon_{ij} = \Upsilon^{auto}(\mathbf{W}_i\mathbf{h}_i)\Upsilon^{auto}(\mathbf{W}_j\mathbf{h}_j)^\top,$$

Table 6: Results of Automatic Neuron Filtering (NeuLT(Auto)).

Embedding	**Model**	**Structure**	**Rest14**		**Laptop14**		**Twitter**	
			Accuracy	*Macro-F1*	*Accuracy*	*Macro-F1*	*Accuracy*	*Macro-F1*
BERT$_{base}$	*Attn.*	Full	85.43	78.04	80.54	77.06	76.22	75.04
	NeuLT(Auto)	Full	**87.04**	**81.19**	**81.96**	**78.70**	**76.96**	**76.01**
RoBERTa$_{base}$	*Attn.*	Full	87.59	81.72	83.86	80.53	75.72	73.92
	NeuLT(Auto)	Full	**88.48**	**83.20**	**84.95**	**82.15**	**78.10**	**77.76**
RoBERTa$_{large}$	*Attn.*	Full	89.46	84.12	84.80	82.19	77.02	75.75
	NeuLT(Auto)	Full	**89.55**	**84.87**	**85.89**	82.19	**78.87**	**78.02**

Neuron-level Analysis

- Method Description

$$\epsilon_{ij} = \Upsilon^{nlt/auto}(\mathbf{h}_i)\Upsilon^{nlt/auto}(\mathbf{h}_j)^\top.$$

Table 8: Neuron statistics in the frequency domain (arranged from high to low frequency).

Pattern	Range	Ratio(%)					
		Rest14		Laptop14		Twitter	
		Train	Test	Train	Test	Train	Test
High	$512 \rightarrow 768$	69.42	69.37	44.36	44.58	69.28	69.19
Bond	$256 \rightarrow 512$	67.93	67.90	42.31	42.15	67.57	67.61
Low	$1 \rightarrow 256$	68.61	69.05	41.93	42.00	67.21	67.18

Table 7: Different frequency filters on RoBERTa$_{base}$. **Bold** indicate improved performance.

Model	Pattern		Rest14(emb)		Laptop14(emb)		Twitter(emb)	
			Acc	F1	Acc	F1	Acc	F1
NeuLT	High	256	87.68	81.28	83.70	80.66	**76.88**	**76.04**
		154	87.50	82.49	83.39	80.22	75.29	74.47
		77	87.59	80.50	83.70	80.99	74.57	72.90
		16	**88.04**	**82.93**	**84.01**	**81.19**	75.29	74.46
		8	**88.13**	**82.09**	83.23	80.43	74.57	74.13
		4	87.32	81.42	**84.33**	**81.69**	74.13	72.50
		2	87.68	81.00	83.86	80.82	74.71	73.13
	Bond	256	87.23	80.83	83.54	80.30	**76.44**	**75.44**
		154	**87.95**	**81.90**	**84.33**	**81.29**	75.58	74.62
		77	87.41	80.64	**84.48**	**81.44**	75.29	74.26
		16	**87.86**	**81.66**	83.70	80.51	75.43	74.45
		8	87.32	81.52	**84.48**	**81.59**	75.14	74.22
		4	**88.04**	**82.19**	83.86	80.89	74.71	73.66
		2	**87.77**	**81.82**	82.92	79.63	75.14	74.00
	Low	256	87.05	80.61	**84.01**	**81.07**	**75.87**	**75.17**
		154	**87.95**	**82.32**	83.54	80.69	75.00	74.07
		77	**88.75**	**82.99**	83.07	79.87	74.28	73.15
		16	**88.84**	**83.33**	83.07	79.22	75.58	74.52
		8	87.32	80.71	**84.33**	**81.71**	74.57	73.46
		4	87.14	80.16	**84.48**	**81.30**	74.86	73.61
		2	87.41	80.96	83.54	80.41	74.71	73.99
Attn.	-		87.77	81.33	84.01	80.94	75.43	74.81
NeuLT(Auto)	-		**87.95**	**81.99**	**84.95**	**82.26**	**76.30**	**75.67**
NeuLT(Auto) w/o DFT	-		87.50	80.54	83.70	80.76	75.87	74.37

Neuron-level Analysis

- Neuron-level Statistics

Table 9: Statistics in the **frequency** domain. **Bold** indicates distinct neurons in top-N.

Dataset		Top-N Neurons (N=10)
Rest14	Train	61, 344, 305, 227, 211, 88, 71, **256**, **168**, **310**
	Test	61, 305, 344, 227, 88, **229**, **32**, **173**, 211, 71
Laptop14	Train	0, 264, 134, 299, 95, 123, 367, **209**, **384**, 281
	Test	0, 264, 134, 367, 299, **192**, 281, 384, 95, 123
Twitter	Train	0, 113, 256, 134, 49, 332, 339, 264, 3, **111**
	Test	0, 113, 256, 264, 49, 332, 3, 134, 339, **115**

Table 10: Statistics in the **time** domain. Underline indicates distinct neurons, and **bold** indicates the same neurons.

Dataset		Top-N Neurons (N=10)
Rest14	Train	**0**, 181, **1**, 594, 304, 269, *675*, 241, 499, 194
	Test	**0**, 181, **1**, 194, 304, 499, 594, 269, *136*, 241
Laptop14	Train	**0**, 32, 689, 642, 383, 747, *280*, 724, *230*, 464
	Test	32, 689, **0**, 383, *626*, 642, *485*, 724, 747, 464
Twitter	Train	**1**, 767, 3, *510*, 691, 645, 116, *344*, *727*, *146*
	Test	**1**, 767, 3, 691, 116, 645, *96*, *448*, *112*, *533*

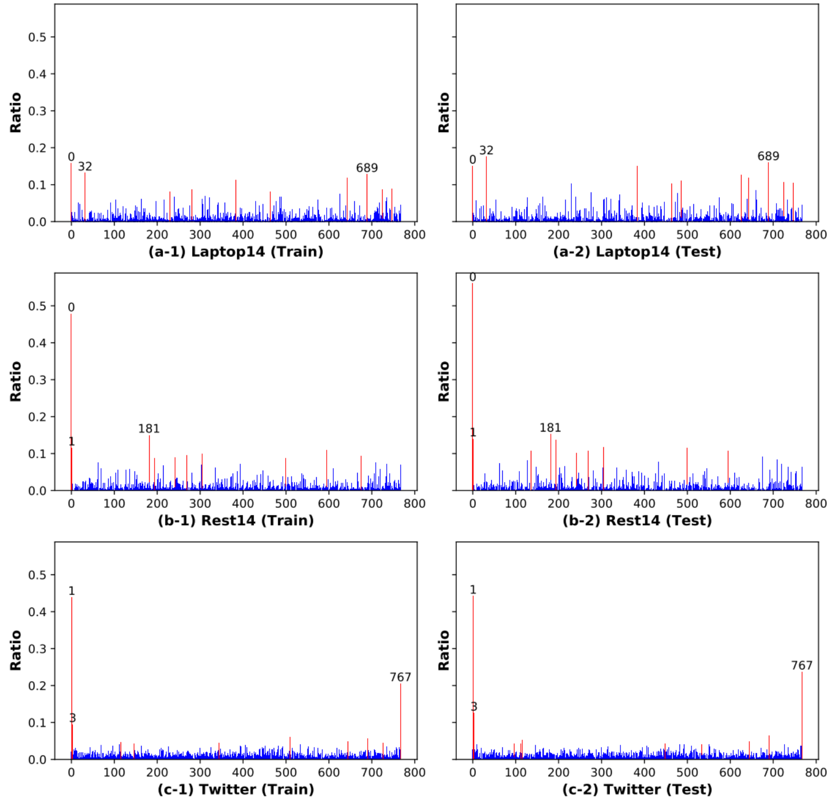

Figure 3: Statistics of Neuron Adjustment. The red line highlights the Top-10 neurons with the highest ratio of adjustment occurrences across all neurons.

Conclusion

- **Conclusion**
  - We propose utilizing GSL to induce latent structures for ABSA by performing a neuron-level manipulation (NeuLT and NeuLT(Auto)) in the frequency domain
  - Extensive experiments and analyses demonstrate that such neuron-level manipulation is effective in structure induction and improvement of ABSA
  - Furthermore, we conducted an in-depth neuron-level analysis to explore this phenomenon.
  - Our exploration is also beneficial to provide inspiration for other similar domains

Thank you for your attention !