Implicit Syntax for Targeted Sentiment Analysis

Yuze Gao, **Yue Zhang** and Tong Xiao

November 27, 2017, TaiWan

Outline

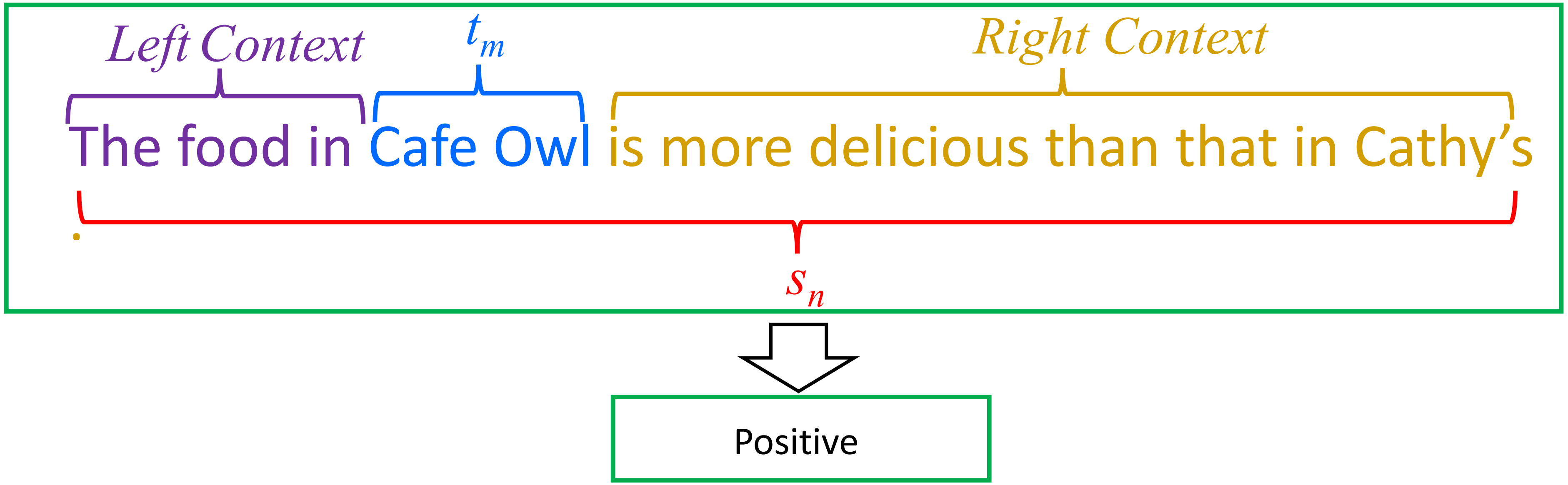
- Background
- Our Method
- Experiments and Results
- Conclusion

Outline

- Background
- Our Method
- Experiments and Results
- Conclusion

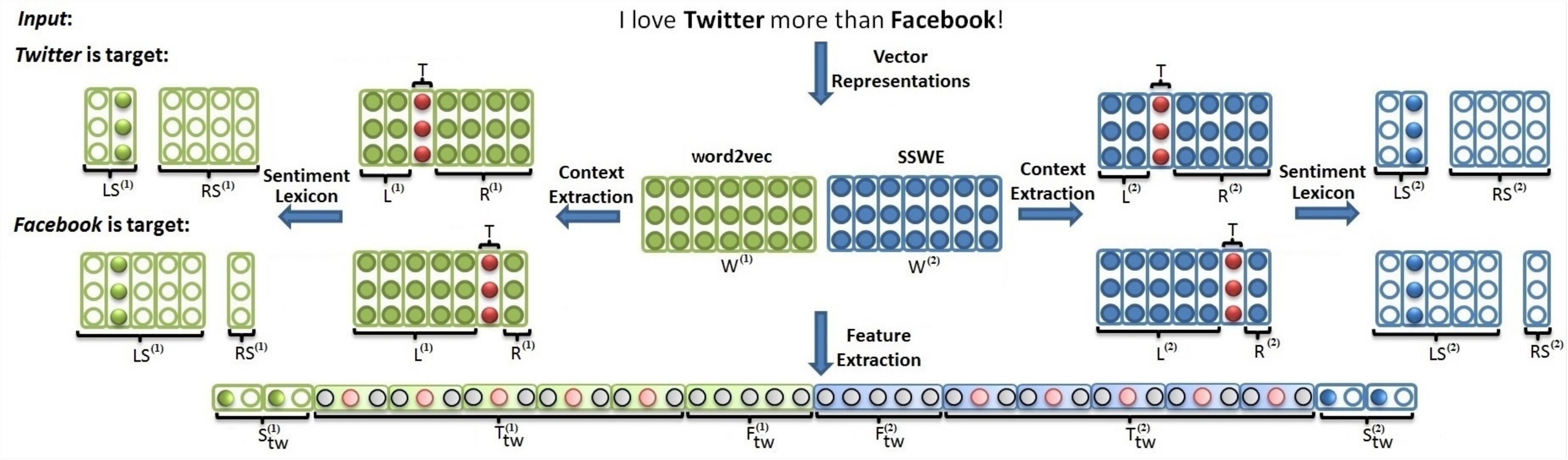
Background

- Targeted Sentiment Analysis
  - Given a sentence s_n with target words t_m, we judge the sentiment(positive, negative, neutral) of the sentence towards the target words t_m.
  - Target Context (Duy-Tin Vo and Yue Zhang 2015)



Background

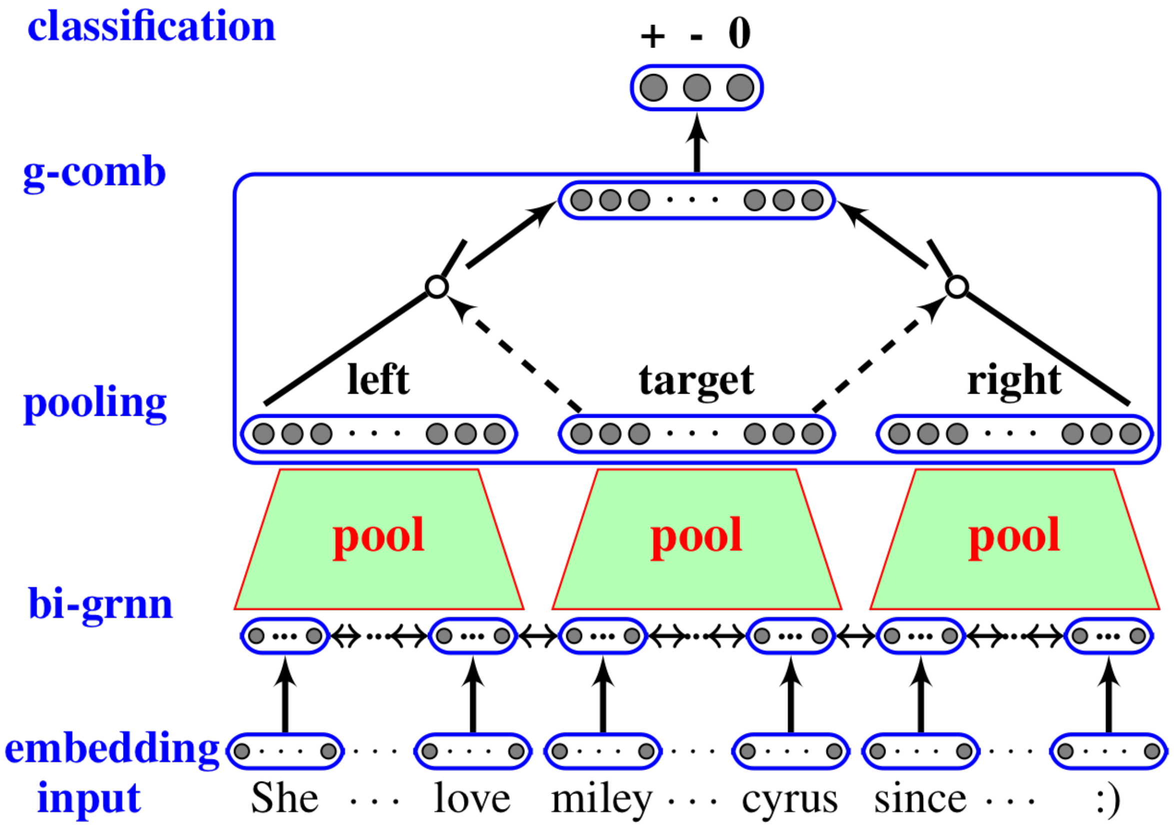
- Baselines:
  - Vo and Zhang (2015)



Duy-Tin Vo and Yue Zhang. 2015. Target-dependent twitter sentiment classification with rich automatic features. In *IJCAI*. pages 1347–1353

Background

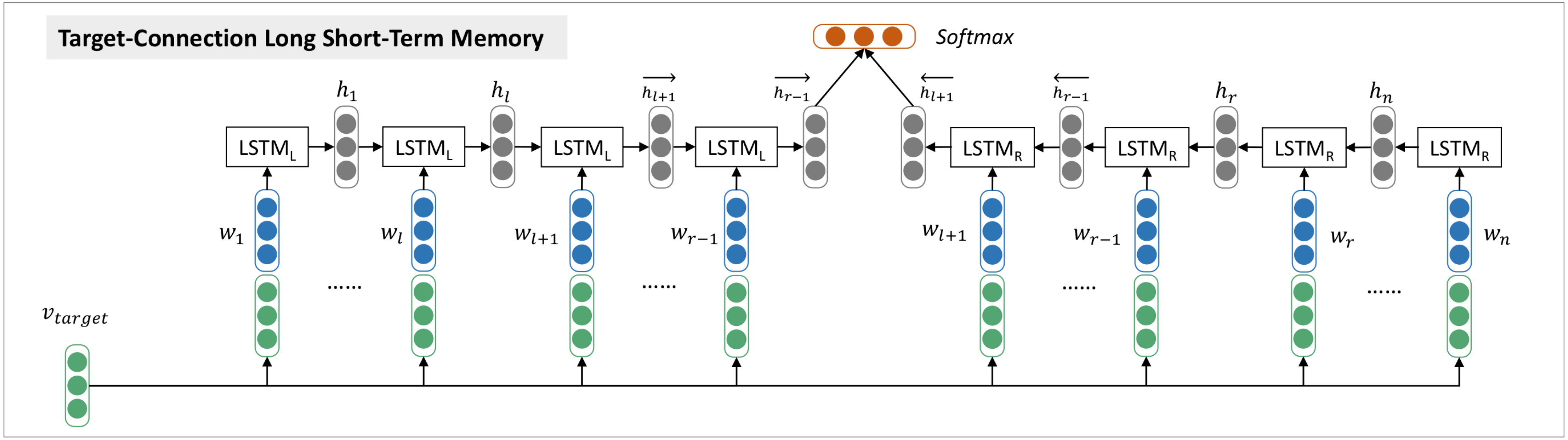
- Baselines:
  - Zhang et al. (2016)



Meishan Zhang, Yue Zhang, and Duy-Tin Vo. 2016. Gated neural networks for targeted sentiment analy- sis. In *AAAI*. pages 3087–3093.

Background

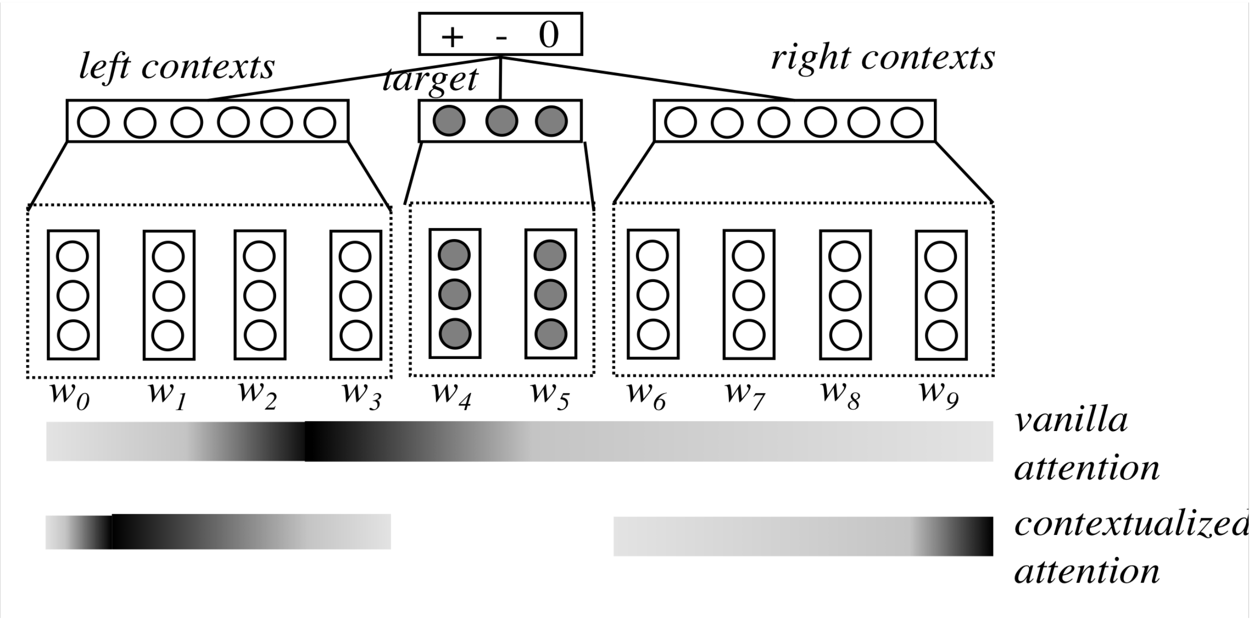
- Baselines:
  - Tang et al. (2016)



Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2016. Effective lstms for target-dependent sentiment classification. In *COLING*. pages 3298–3307.

Background

- Baselines:
  - Liu and Zhang (2017)



Jiangming Liu and Yue Zhang. 2017. Attention mod- eling for targeted sentiment. *EACL 2017* page 572.

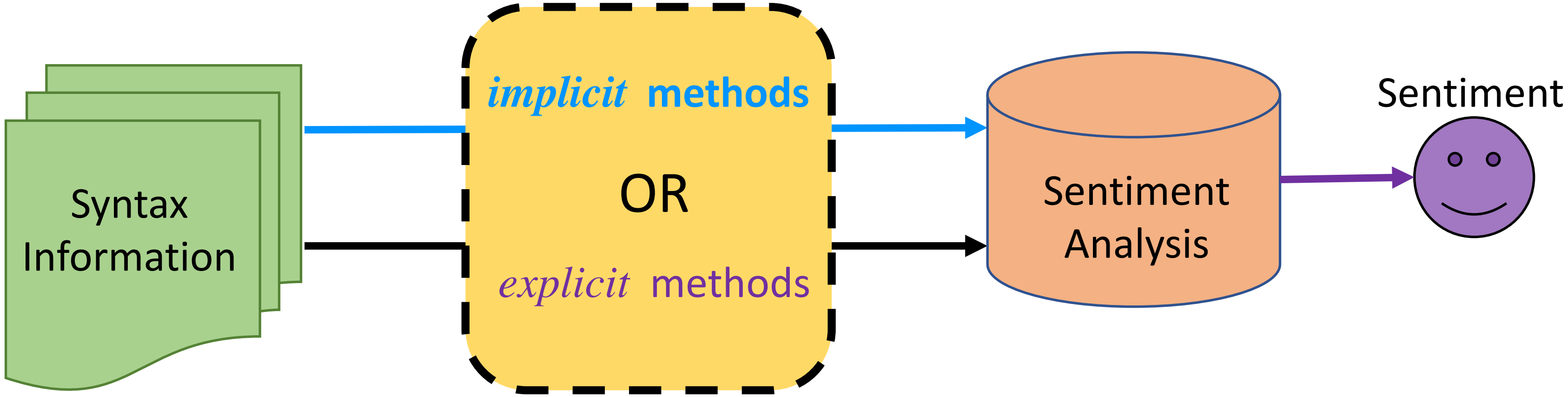
Background

- Baselines:
  - Results

	Acc.(%)		F1(%)	
Models	Z_{set}	T_{set}	Z_{set}	T_{set}
Vo and Zhang (2015)	69.6	71.1	65.6	69.9
Tang et al. (2015)	/	71.5	/	69.5
Zhang et al. (2016)	71.9	72.0	69.6	70.9
Liu and Zhang (2017)	73.5	72.4	70.6	70.5

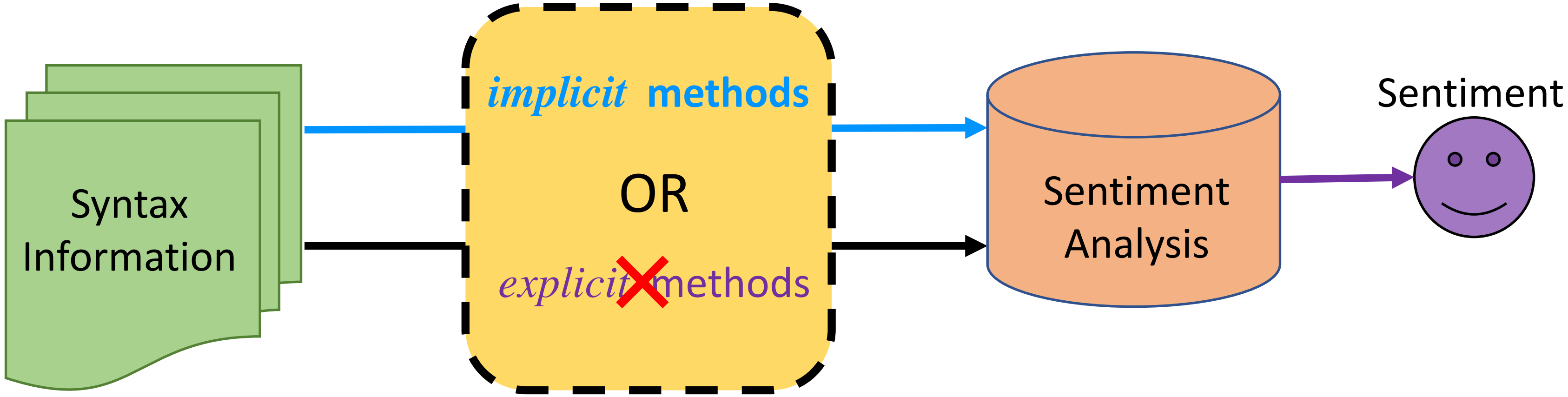
Background

- Syntax information should help improve targeted sentiment analysis
- Purpose: how to utilize the syntax information properly



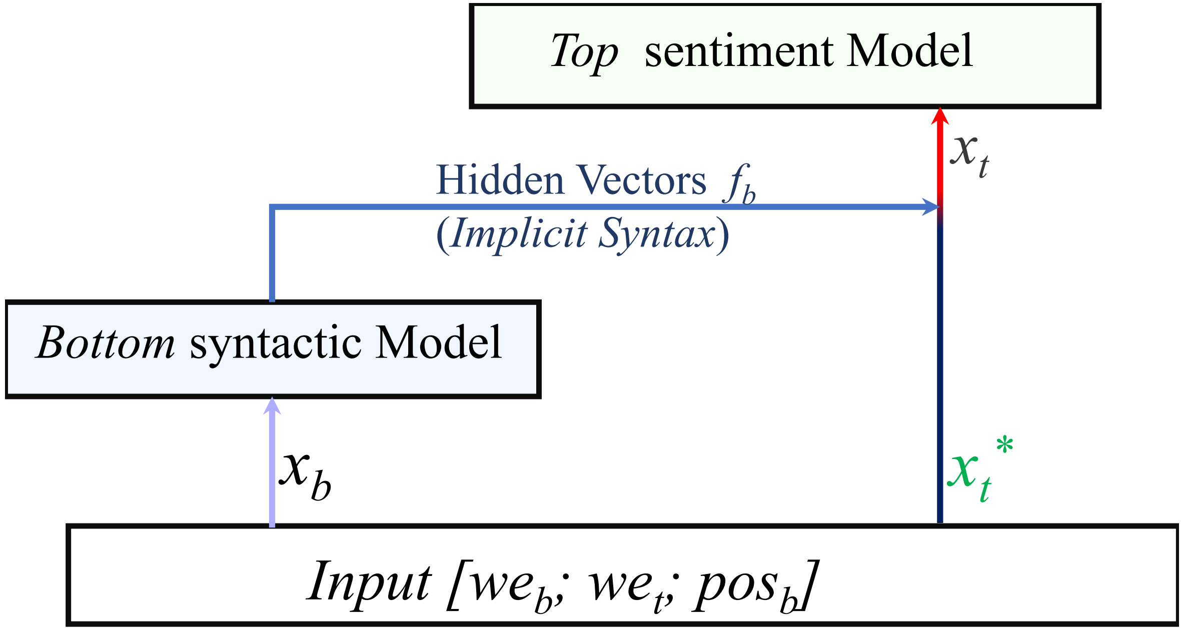
Background

- Directly using parser output suffer noise and error propagation



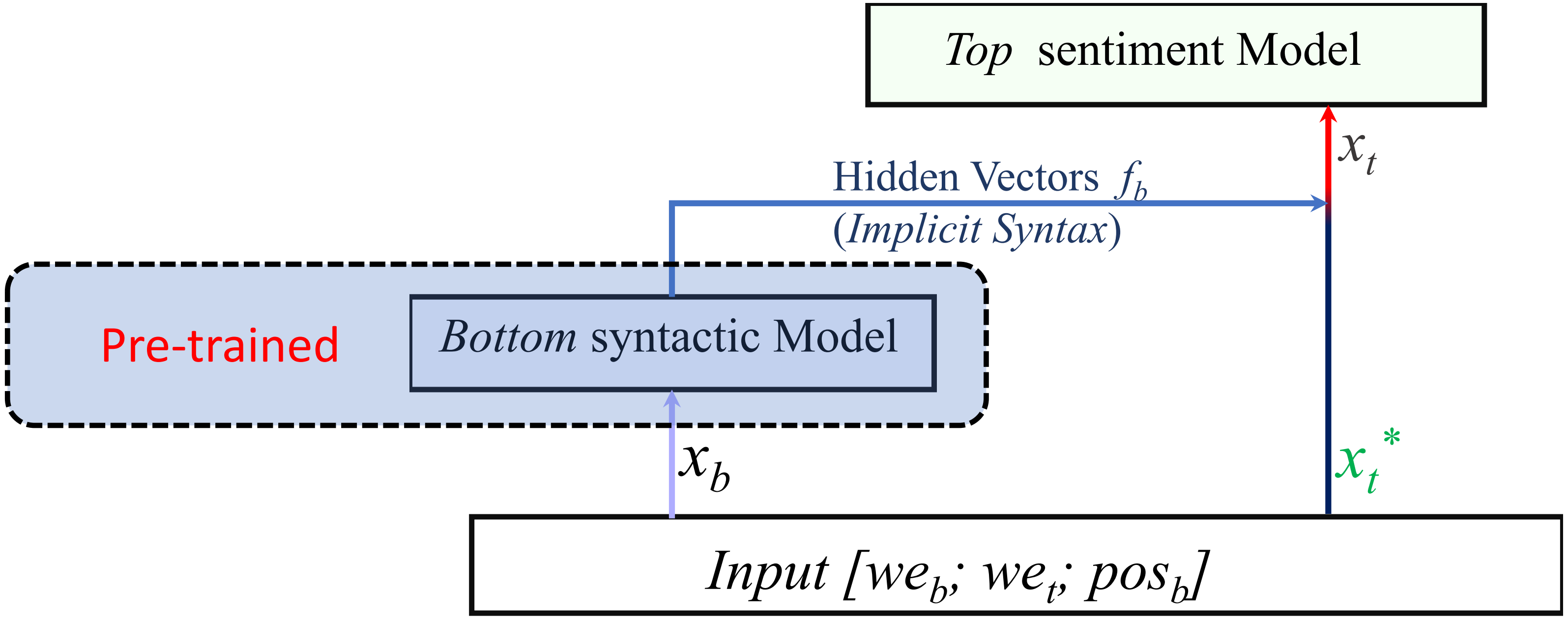
Our Method

- Model Structure



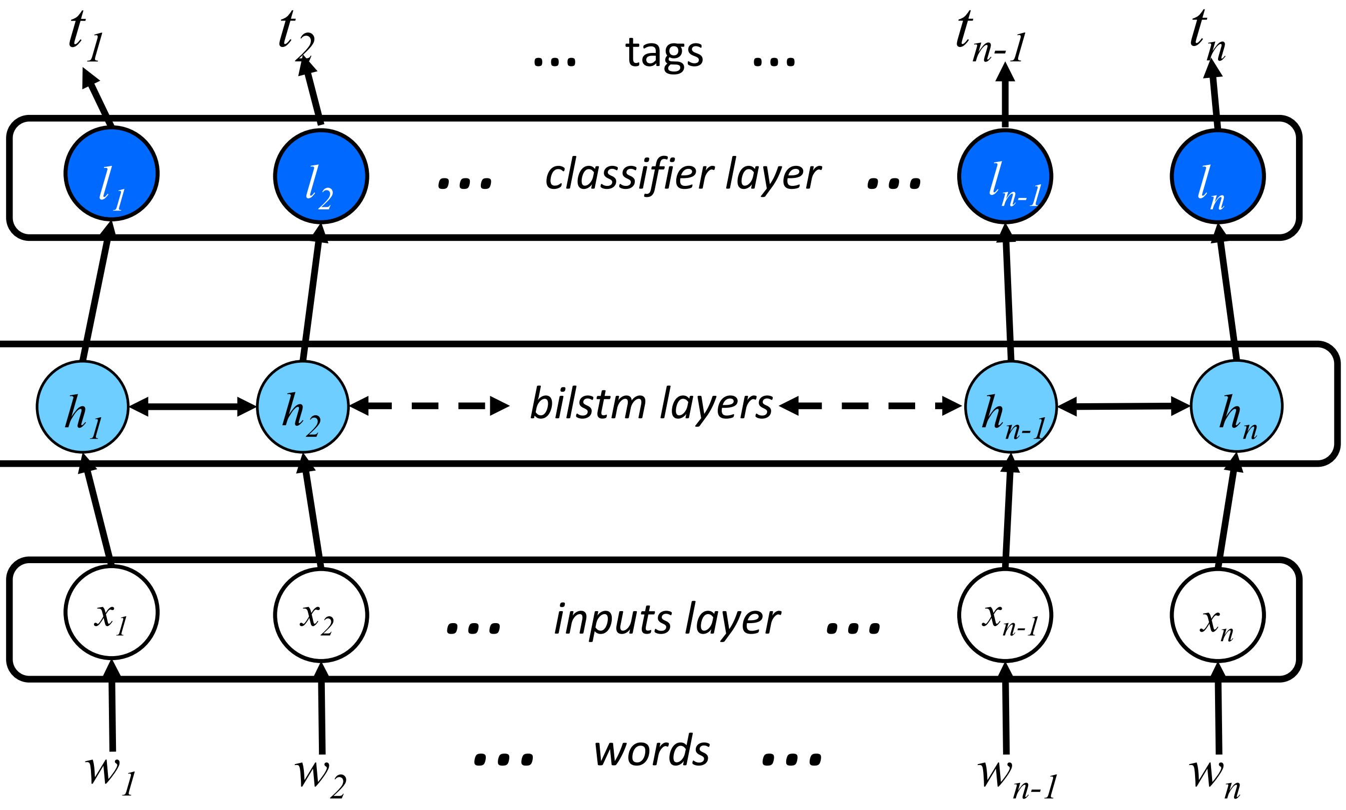
Our Method

- Model Structure



Our Method

- POS-tagging model

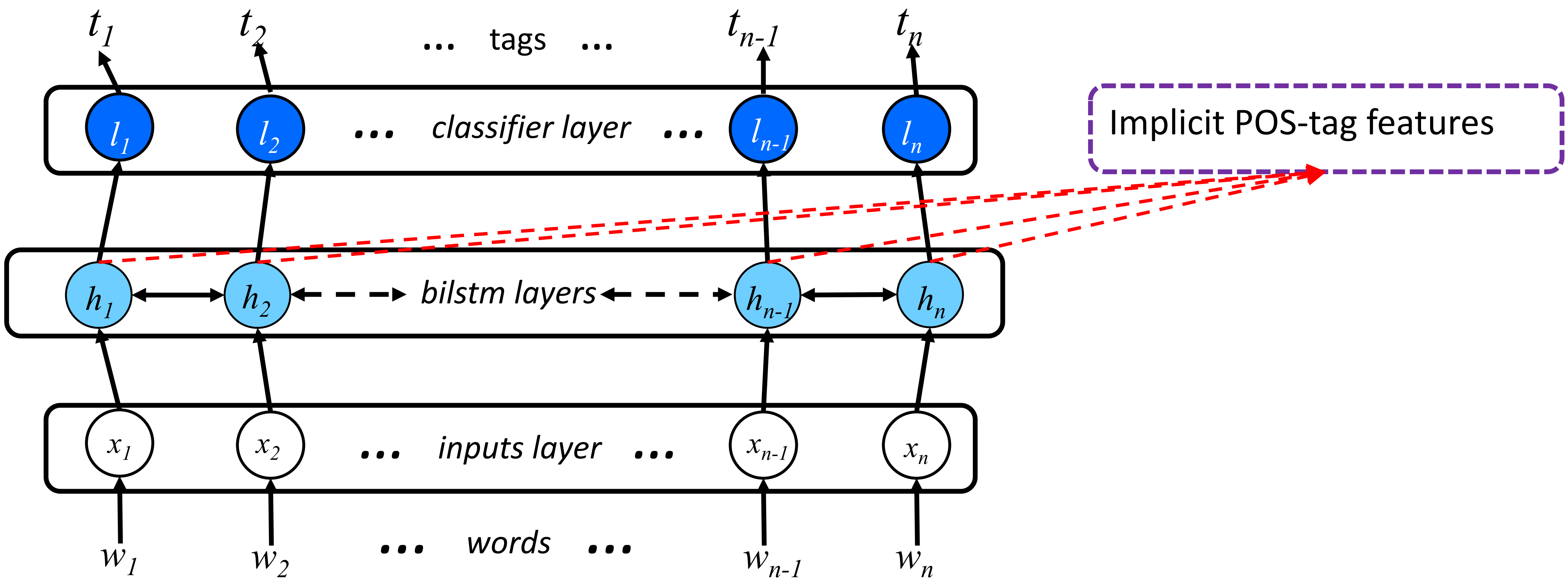


$$Labels = \mathrm{Classifier}(W_1 S' + b)$$

$$\begin{aligned} S' &= [h_1, h_2, ..., h_n] \\ &= \mathrm{BiLSTM}([x_1, x_2, ..., x_n])^{k_1} \end{aligned}$$

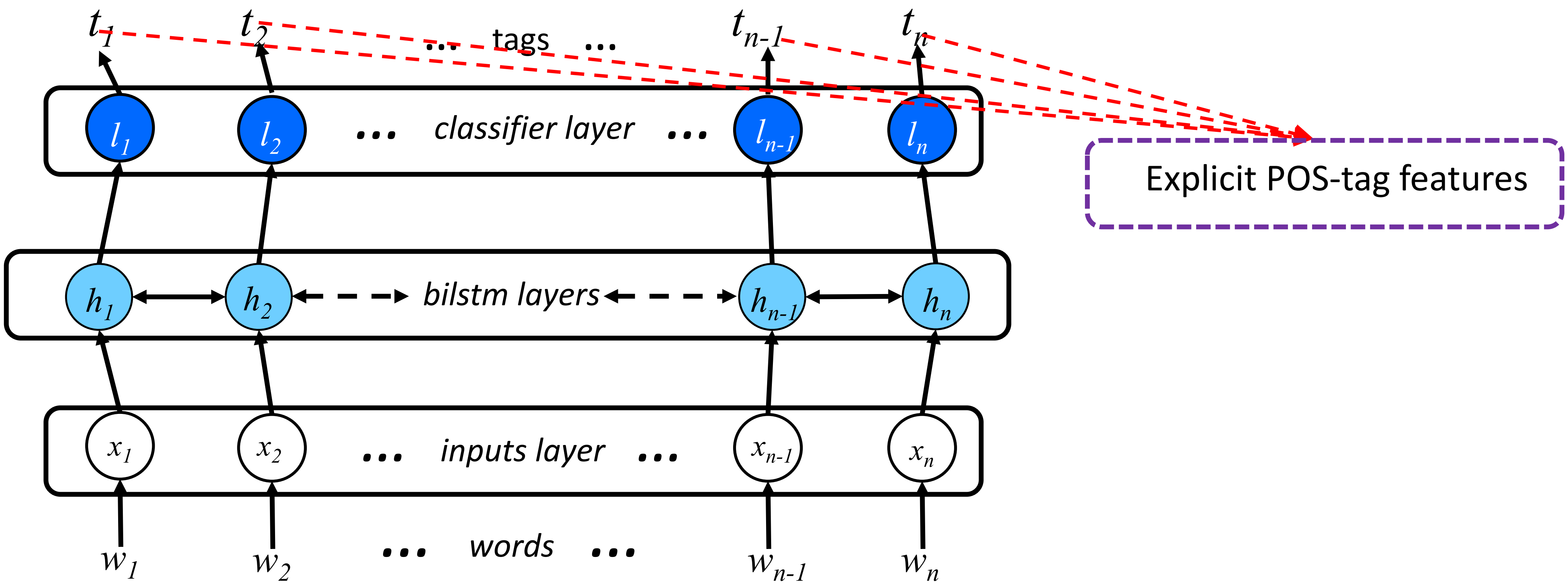
Our Method

- POS-tagging model



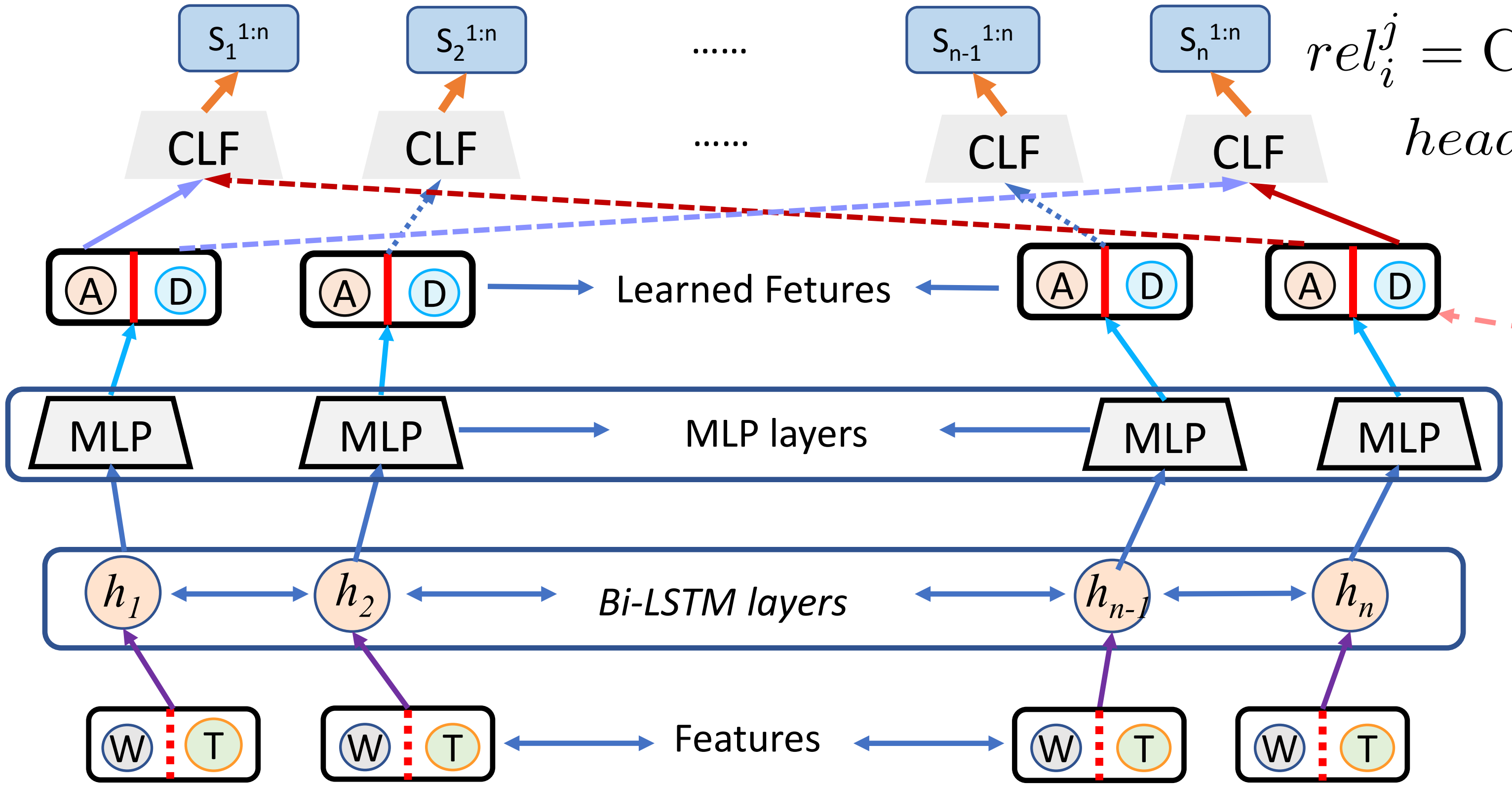
Our Method

- POS-tagging model



Our Method

- Dependency model



$$rel_i^j = \text{Classifier}(head_{arc}^i, head_{dep}^j, head_i^j)$$

$$head_i^j = \text{Classifier}(head_{arc}^i, head_{dep}^j)$$

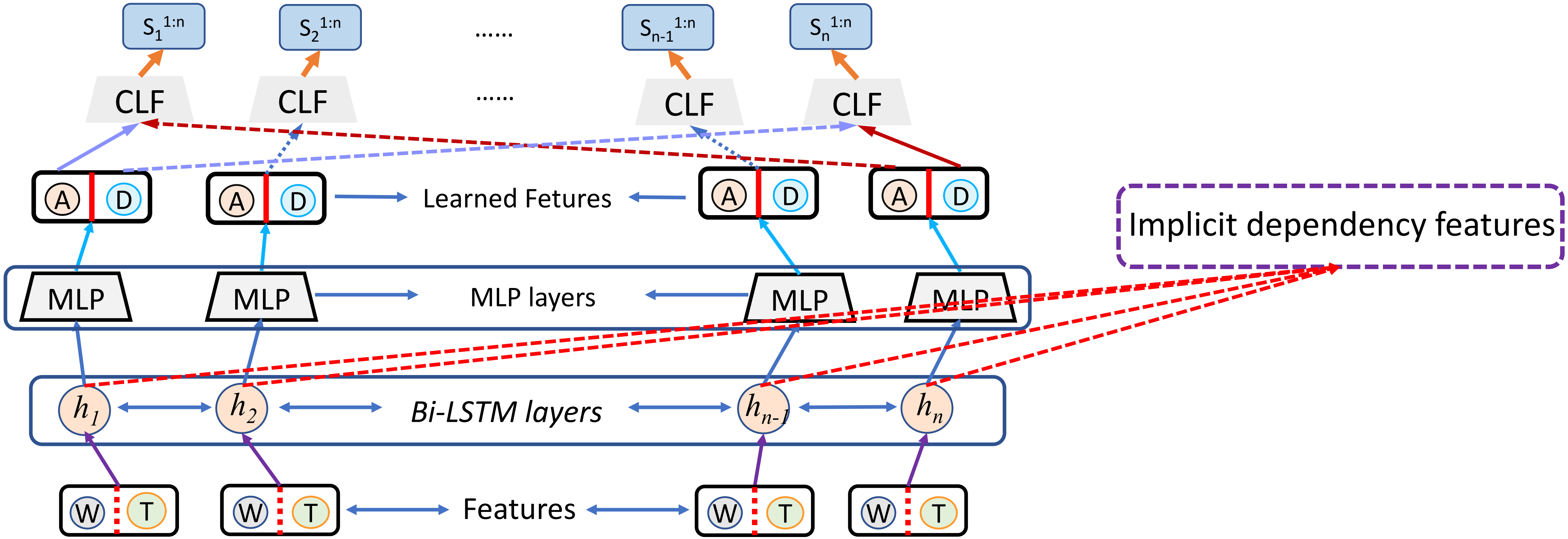
$$\boxed{head_{arc}, head_{dep}}, \boxed{rel_{arc}, rel_{dep}} = \text{MLP}([h_1, h_2, ..., h_n])^{k_3},$$

$$\begin{aligned} S' &= [h_1, h_2, ..., h_n] \\ &= \text{BiLSTM}([x_1, x_2, ..., x_n])^{k_2} \end{aligned}$$

Timothy Dozat and Christopher D Manning. 2016. Deep biaffine attention for neural dependency pars- ing. *arXiv preprint arXiv:1611.01734* .

Our Method

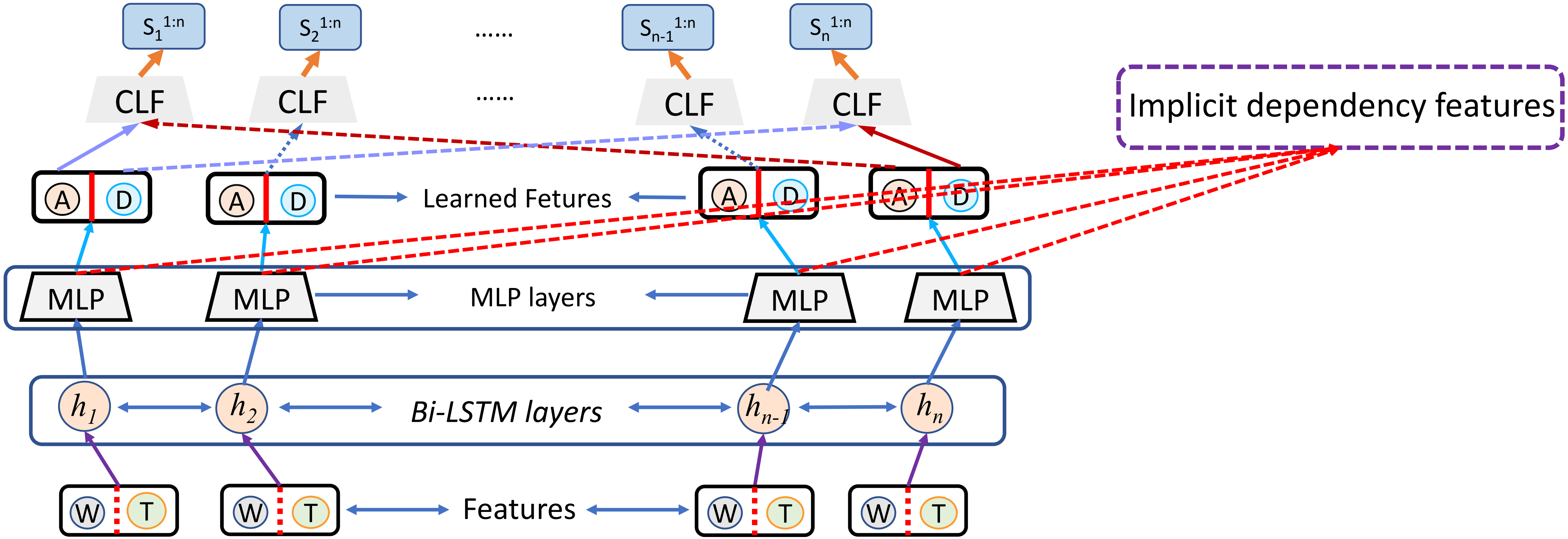
- Dependency model



Timothy Dozat and Christopher D Manning. 2016. Deep biaffine attention for neural dependency pars- ing. *arXiv preprint arXiv:1611.01734* .

Our Method

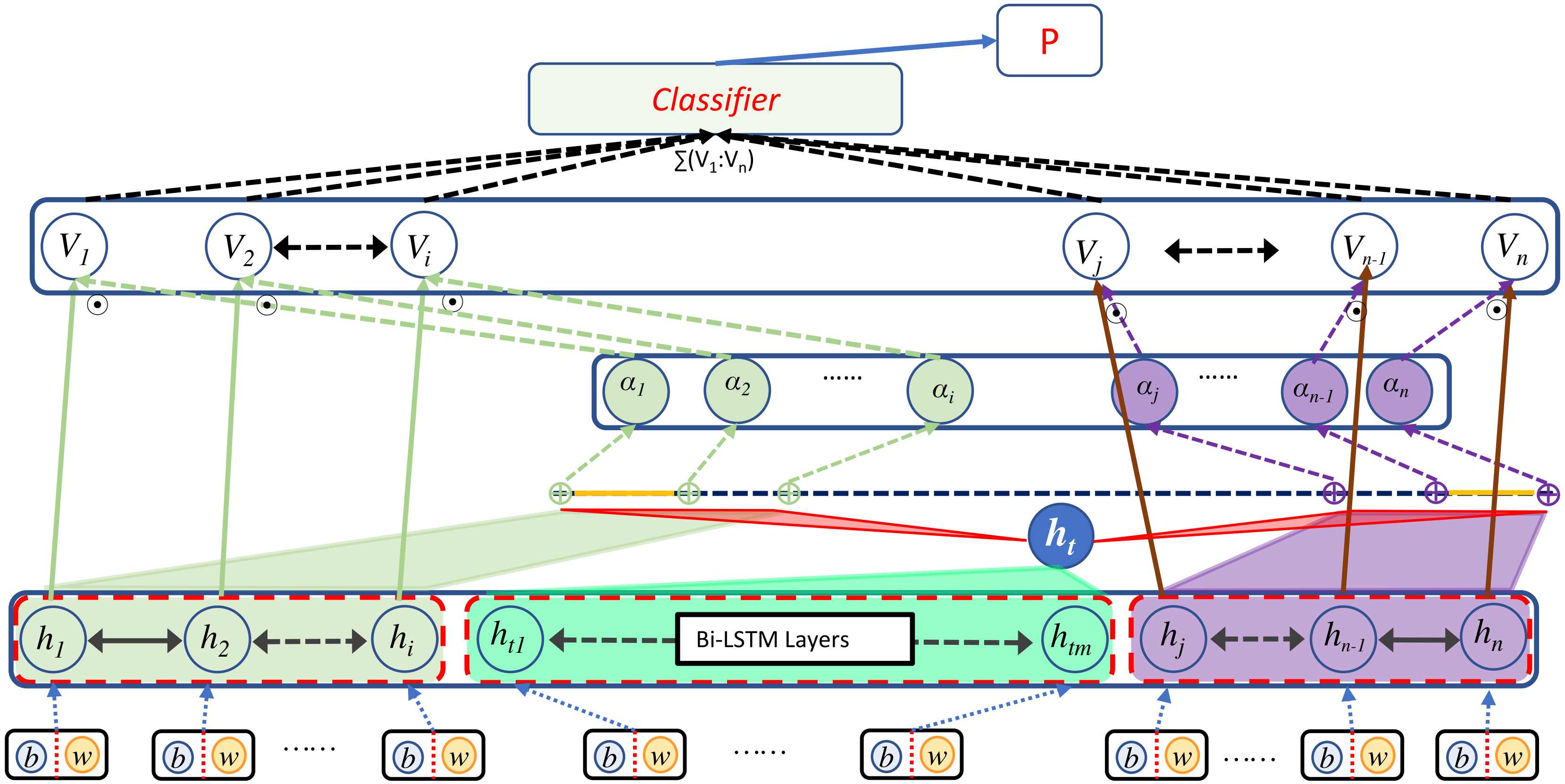
- Dependency model



Timothy Dozat and Christopher D Manning. 2016. Deep biaffine attention for neural dependency pars- ing. *arXiv preprint arXiv:1611.01734* .

Our Method

- Top Sentiment Model



$$P = \text{Classifier}(W_3 \cdot S'_\alpha + b_2)$$

$$S'_\alpha = \text{Attention}([h_1, h_2, ..., h_n], h_t) = \sum_{i=1}^{n} \alpha_i h_i = \sum_{i=1}^{n} V_i$$

$$h_t = \frac{1}{m} \sum_{i=1}^{m} h_{t_i}$$

$$[h_1, h_2, ..., h_n] = \text{BiLSTM}([r_1, r_2, ..., r_n])^{k_4}$$

Implicit Features and Word-embeddings

Outline

- Background
- Our Method
- Experiments and Results
- Conclusion

Experiments and Results

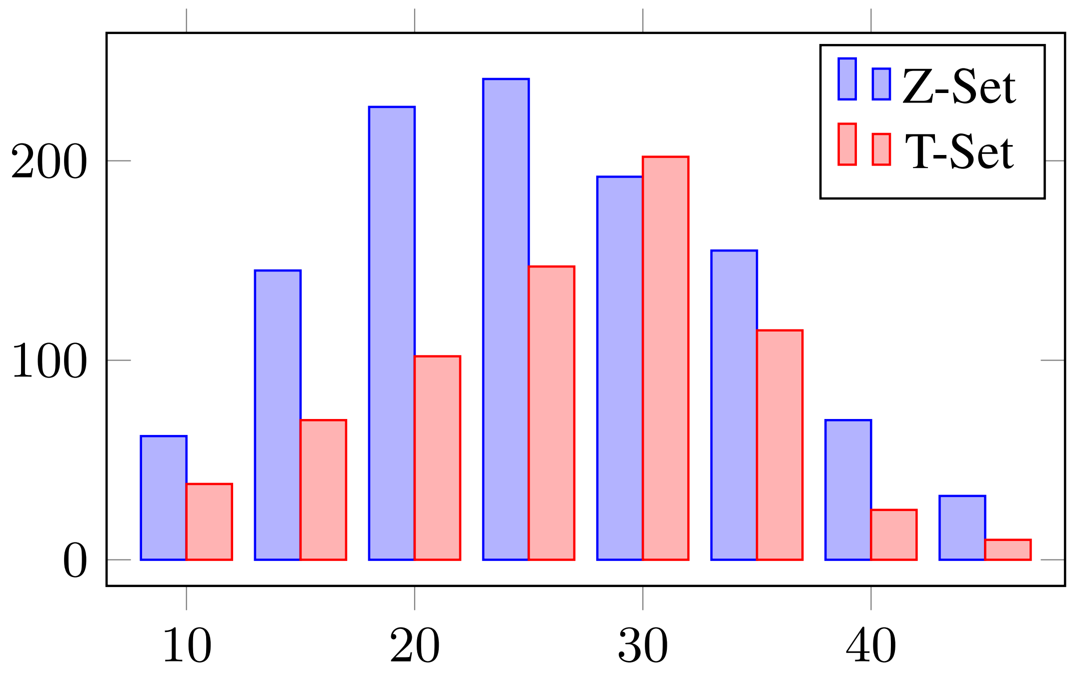
- Dataset
  - PTB3 [standard splits]
  - Z-Set [Zhang et al. (2016)]
  - T-Set [Tang et al. (2015)]

		Total	Pos	Neg	Neu
T-set	Train	6248	1561	1560	3127
	Test	692	173	173	346
Z-set	Train	9489	2416	2384	4689
	Dev	1036	255	272	509
	Test	1170	294	295	581

Meishan Zhang, Yue Zhang, and Duy-Tin Vo. 2016. Gated neural networks for targeted sentiment analysis. In *AAAI*. pages 3087–3093.
Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2015. Effective lstms for target-dependent sentiment classification. In *COLING*. pages 3298–3307.

Experiments and Results

- Dataset
  - Test Set sentence length distribution



Experiments and Results

- The Results of Bottom syntactic Model(Pre-trained part)

Models	Acc.(%)	F1(%)	UAS	LAS
POS-tagging	92.4	91.6	/	/
Normal Dep.	/	/	95.6	93.8
No-POS Dep.	/	/	94.3	92.7

Experiments and Results

- Results of test set with different implicit syntax features

	Acc.(%)		F1(%)	
Models	Z_{set}	T_{set}	Z_{set}	T_{set}
Baseline	*73.0*	*71.7*	*70.2*	*70.1*
$+\ lm_{pos}$ [a]	73.5	72.4	71.2	70.4
$+\ lt_{pos}$ [b]	73.2	72.0	70.8	70.2
$+\ lm_{pos}\&lt_{pos}$ [c]	**73.9**	**72.5**	**71.4**	**70.7**
$+\ lm_{dep}$	73.5	72.2	70.7	70.6
$+\ mlp_{dep}$	74.0	72.6	71.3	70.9
$+\ lm_{dep}\&mlp_{dep}$	**74.1**	**72.7**	**71.7**	**71.3**
$+\ lm^{*}_{dep}$ [d]	73.3	72.4	70.9	70.5
$+\ mlp^{*}_{dep}$ [e]	74.2	72.8	71.3	70.5
$+\ lm^{*}_{dep}\&mlp^{*}_{dep}$ [f]	**74.3**	**72.8**	**71.8**	**71.4**

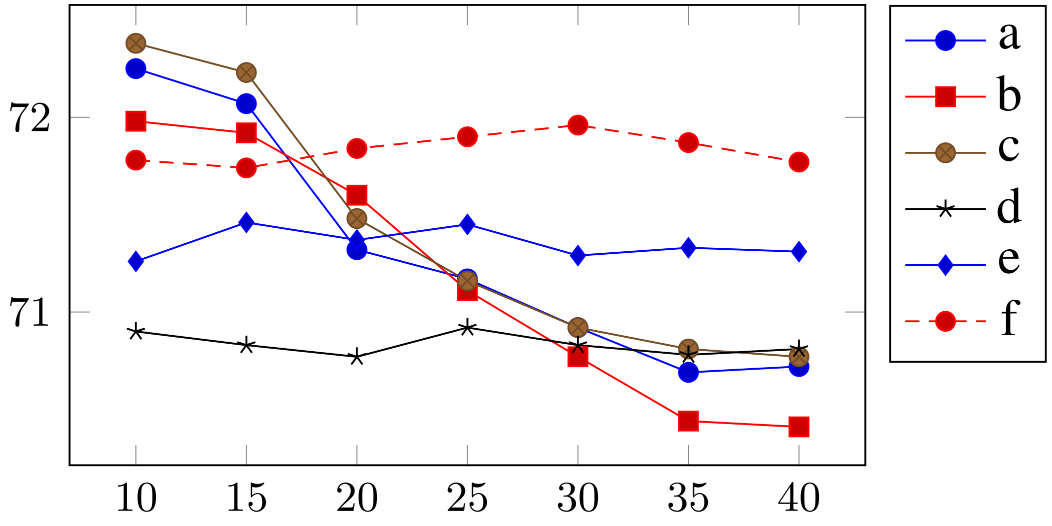
Experiments and Results

- F1 values of each polarity on test set

		Pos	Neg	Neu
Z-Set	Baseline	61.64	69.83	78.67
	POS	61.43	70.17	78.97
	$DEP_{[2]}$	61.14	71.14	79.63
T-Set	Baseline	62.57	69.36	75.70
	POS	61.84	69.41	77.62
	$DEP_{[2]}$	62.74	70.31	78.42

Experiments and Results

- Test set F1 values against sentence length(Z-Dataset)



Conclusion

- Implicit syntax features by neural stacking method can obviously help enhance the targeted sentiment analysis.
- POS-tagging features can carry more positive implicit information that help short sentence.
- Dependency implicit features show robust and stable in different sentence length.

Thank you!

Code is available at *https://github.com/CooDL/TSSSF*