Sentiment Classification of Chinese Contrast Sentences

Junjie Li[1], Yu Zhou[1], Lin Pang[2], Chunyang Liu[2]

[1] National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences

[2] National Computer Network Emergency Response Technical Team/Coordination Center of China

Outline

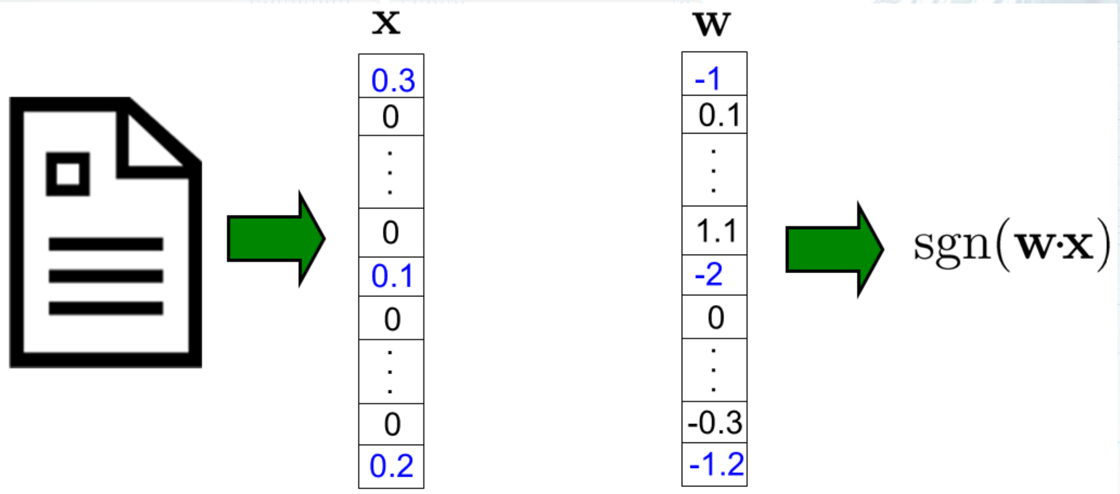
- **Motivation**
- Related Work
- Approach Overview
- Methodology
- Experiments
- Conclusions

Sentence-level Sentiment Classification

- Input: a sentence
- Output: sentiment polarity of the sentence
  - Positive/Negative

Sentence-level Sentiment Classification

- Input: a sentence
- Output: sentiment polarity of the sentence
  - Positive/Negative
- Method: bag-of-words(BOW) + Classification



Drawbacks

- BOW ignores the relationship among words and sentences.

Drawbacks

- BOW ignores the relationship among words and sentences.
- Useful linguistic phenomena hard to handle.

Drawbacks

- BOW ignores the relationship among words and sentences.
- Useful linguistic phenomena hard to handle.
  - Negation: 我不喜欢这家酒店 (I don’t like the hotel).
  - Contrast: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

How to Improve

- For relationship among words:
  - n-Gram, dependency

How to Improve

- For relationship among words:
  - n-Gram, dependency
- For relationship among sentences:
  - Discourse structure
  - Compound Sentence structure

How to Improve

- For relationship among words:
  - n-Gram, dependency
- For relationship among sentences:
  - Discourse structure
  - Compound Sentence structure

Help

Chinese Contrast Sentences Sentiment classification

How to Improve

- For relationship among words:
  - n-Gram, dependency
- For relationship among sentences:
  - Discourse structure
  - Compound Sentence structure

Help

We focused on

Chinese Contrast Sentences Sentiment classification

Contrast Sentence Structure

- Two clauses in contrast sentence:
  - nuclei clause
  - satellite clause

Contrast Sentence Structure

- Two clauses in contrast sentence:
  - nuclei clause
  - satellite clause

酒店位置好 (The hotel location is good)	但是设施糟糕 (but its facilities are poor)
satellite clause	nuclei clause

Outline

- Motivation
- **Related Work**
- Approach Overview
- Methodology
- Experiments
- Conclusions

Related Work

The hotel location is good, but its facilities are poor

Word-level weight w_1 w_2 w_3 w_4 w_5 w_6 w_7

Clause-level weight $w_{(s1)}$ $w_{s(2)}$

Related Work

The hotel location is good, but its facilities are poor

Word-level weight w_1 w_2 w_3 w_4 w_5 w_6 w_7

Clause-level weight $w_{(s1)}$ $w_{s(2)}$

- SNSS (Single Nucleus Single Satellite Method)[2,11,13] gets clause-level weights according to different clauses in contrast relation.

Related Work

The hotel location is good, but its facilities are poor

Word-level weight w_1 w_2 w_3 w_4 w_5 w_6 w_7

Clause-level weight $w_{(s1)}$ $w_{s(2)}$

- SNSS (Single Nucleus Single Satellite Method)[2,11,13] gets clause-level weights according to different clauses in contrast relation.
- Joint strategy[5]: simultaneity learns both two BOW model (satellite part and nuclei part).

We Can Do Better……

- Q: Different connectives result in different label.
  - 酒店位置好(The hotel location is good)，只是设施糟糕 (yet its facilities are poor)
  - 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

We Can Do Better……

- Q: Different connectives result in different label.
  - 酒店位置好(The hotel location is good)，只是设施糟糕 (yet its facilities are poor)
  - 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)
- A: Consider connectives not just contrast relation.

We Can Do Better……

- Q: Different connectives result in different label.
  - 酒店位置好(The hotel location is good)，只是设施糟糕 (yet its facilities are poor)
  - 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)
- A: Consider connectives not just contrast relation.
- Q: The pipeline approach often causes error propagation.
- A: Propose a model which jointly learn weights in both clause-level and word-level together.

We Can Do Better……

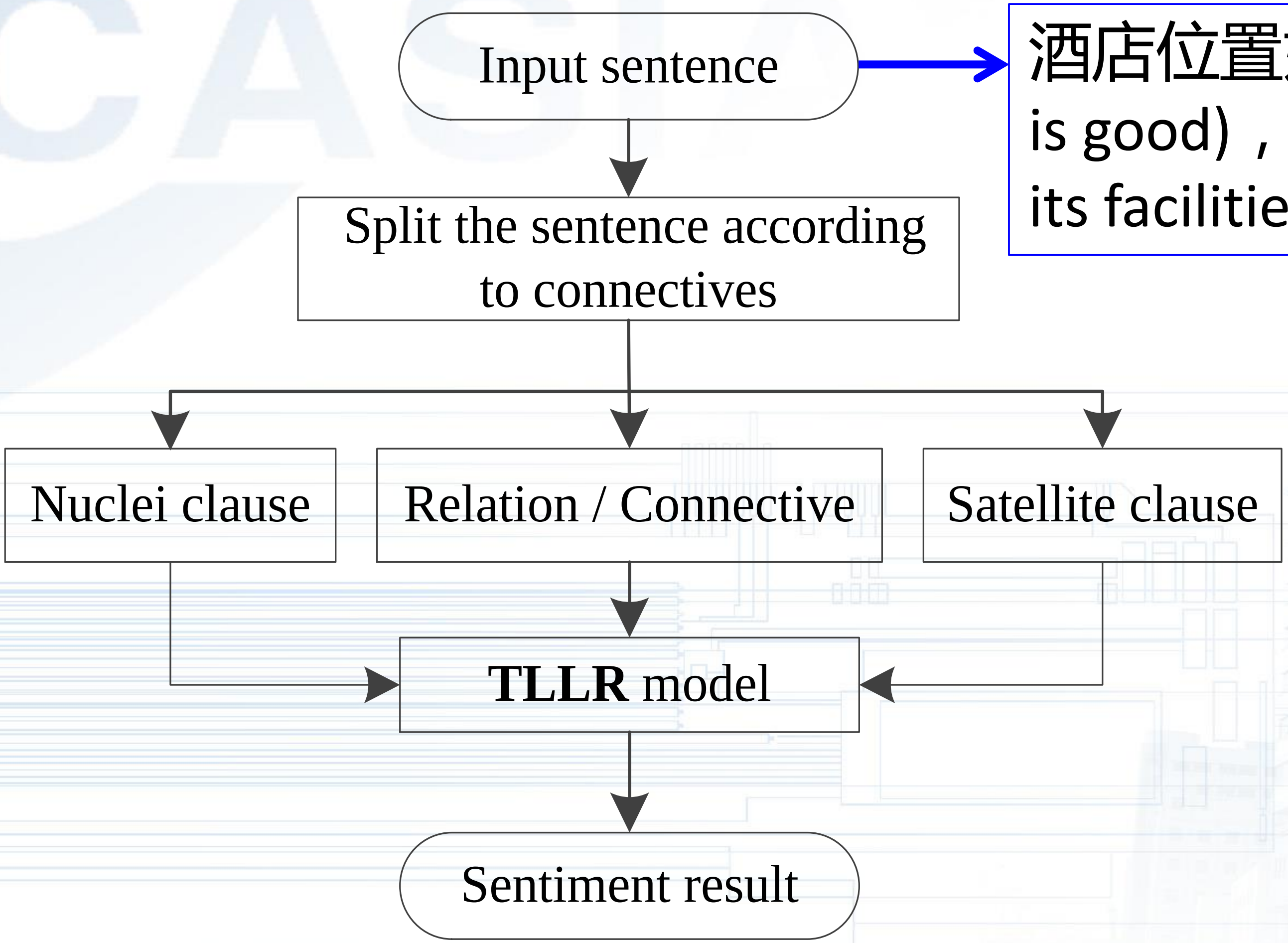
- Q: Different connectives result in different label.
  - 酒店位置好(The hotel location is good)，只是设施糟糕 (yet its facilities are poor)
  - 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)
- A: Consider connectives not just contrast relation.
- Q: The pipeline approach often causes error propagation.
- A: Propose a model which jointly learn weights in both clause-level and word-level together.

Two-Layer Logistic Regression (TLLR) model

Outline

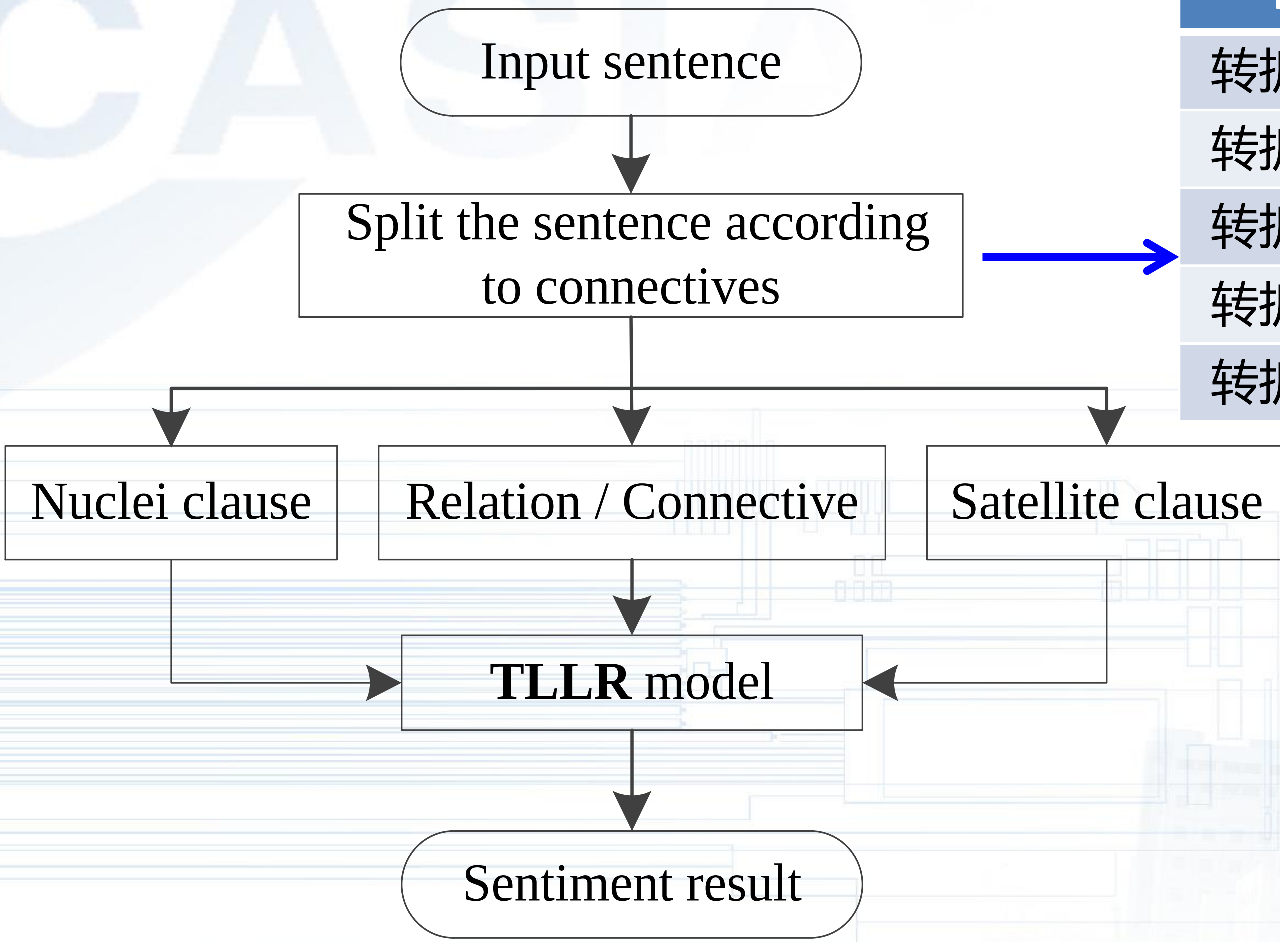
- Motivation
- Related Work
- **Approach Overview**
- Methodology
- Experiments
- Conclusions

Approach Overview



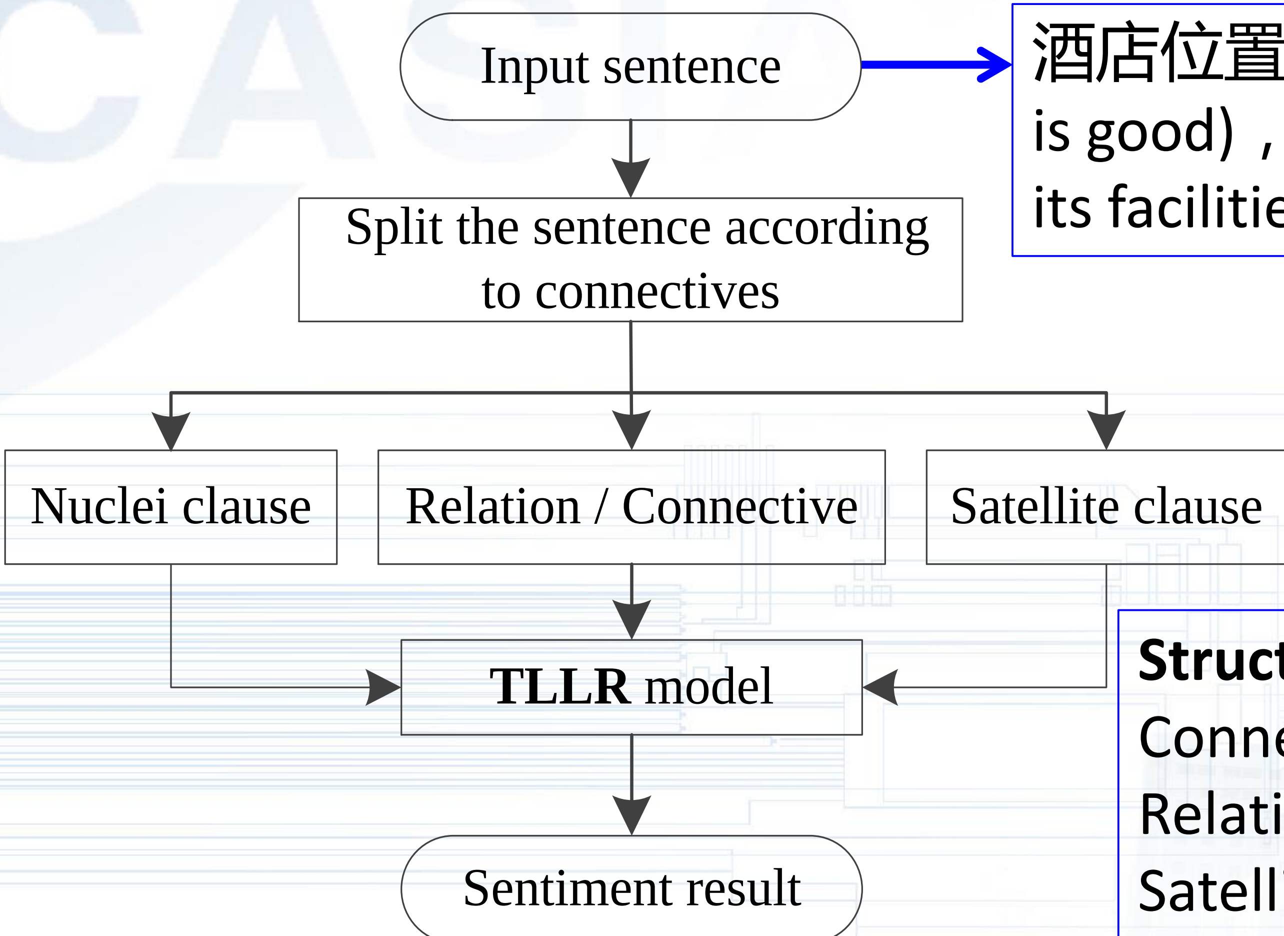
酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Approach Overview



Relation	Connectives
转折(Contrast)	不过(however)
转折(Contrast)	但是(but)
转折(Contrast)	但(but)
转折(Contrast)	只是(yet)
转折(Contrast)	可是(however)

Approach Overview



酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Structure Output:

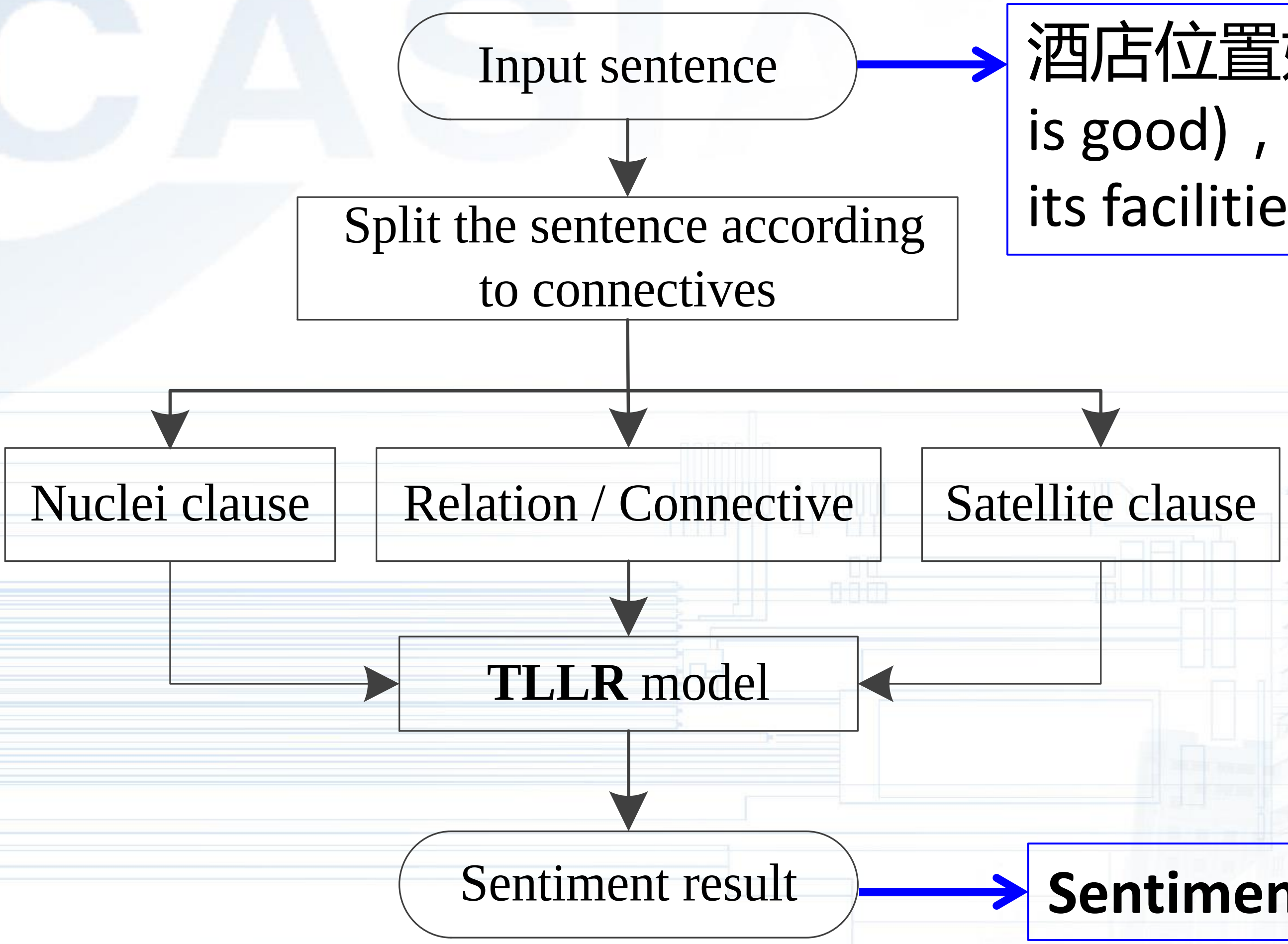
Connective: 但是

Relation: 转折

Satellite Part: 酒店位置好

Nuclei Part: 但是设施糟糕

Approach Overview



酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Sentiment Output: Negative

Outline

- Motivation
- Related Work
- Approach Overview
- **Methodology**
  - Two-Layer Logistic Regression (TLLR) model
- Experiments
- Conclusions

Logistic Regression

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Logistic Regression

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Feature Vector($\vec{x}$): 酒店 位置 好 ， 但是 设施 糟糕

Word Weight Vector($\vec{\theta}$): θ(酒店) θ(位置) θ(但是) θ(糟糕)

Logistic Regression

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Feature Vector($\vec{x}$): 酒店 位置 好 ， 但是 设施 糟糕

Word Weight Vector($\vec{\theta}$): θ(酒店) θ(位置) θ(但是) θ(糟糕)

Classification: $$P(y=1|\vec{x}) = h(\vec{\theta}\cdot\vec{x}) = \frac{1}{1+e^{-\vec{\theta}\cdot\vec{x}}}$$

Logistic Regression

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

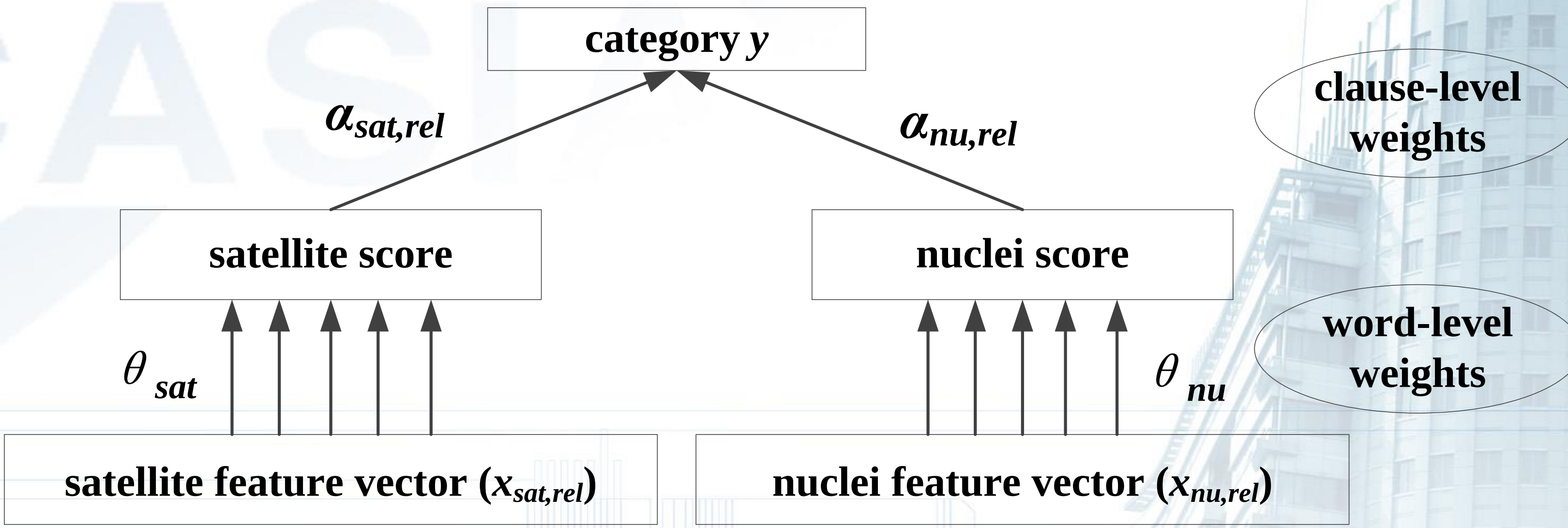
Feature Vector($\vec{x}$): 酒店 位置 好 ， 但是 设施 糟糕

Word Weight Vector($\vec{\theta}$): θ(酒店) θ(位置) θ(但是) θ(糟糕)

Classification: $$P(y=1|\vec{x}) = h(\vec{\theta}\cdot\vec{x}) = \frac{1}{1+e^{-\vec{\theta}\cdot\vec{x}}}$$

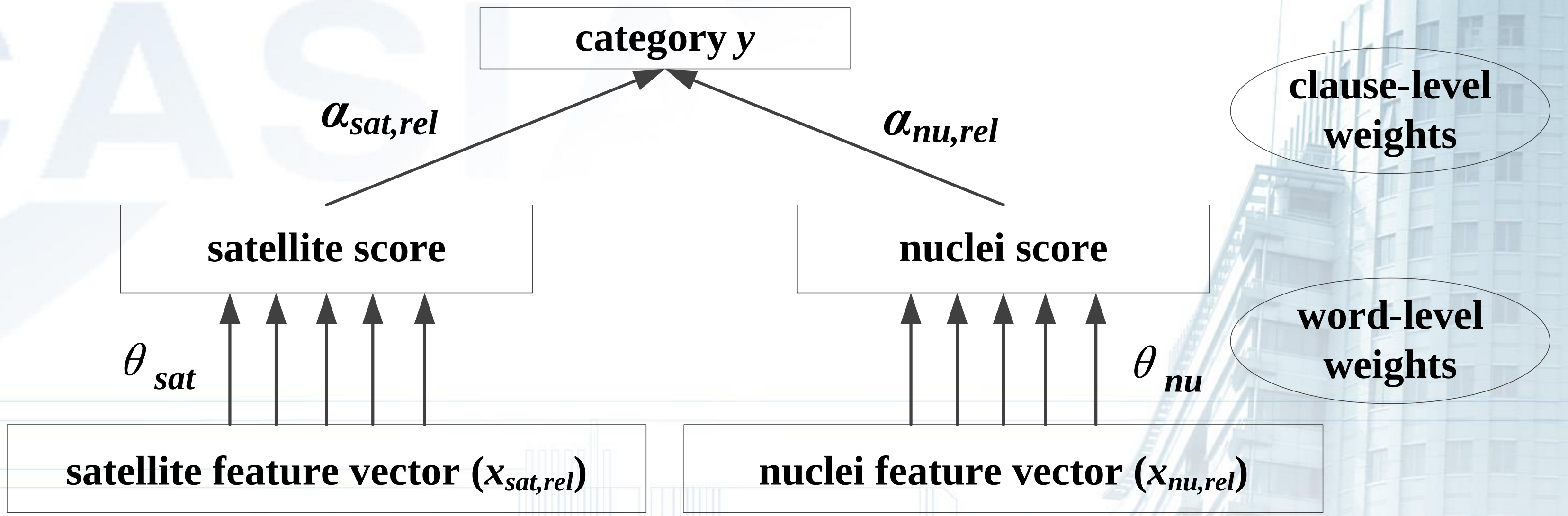
Loss Function: $$Cost(h(\vec{\theta}\cdot\overrightarrow{x_k}), y_k) = \begin{cases} -log(h(\vec{\theta}\cdot\overrightarrow{x_k})) & if\ \ y_k = 1 \\ -log(1-h(\vec{\theta}\cdot\overrightarrow{x_k})) & if\ \ y_k = 0 \end{cases}$$

Two-Layer Logistic Regression



$$P(y=1\mid\vec{x}) = h\left(\alpha_{sat,rel}\times(\overline{\theta_{sat}}\cdot\overline{x_{sat,rel}})+\alpha_{nu,rel}\times(\overline{\theta_{nu}}\cdot\overline{x_{nu,rel}})\right)$$

Two-Layer Logistic Regression



$$P(y=1|\vec{x}) = h\left(\alpha_{sat,rel} \times (\overrightarrow{\theta_{sat}} \cdot \overrightarrow{x_{sat,rel}}) + \alpha_{nu,rel} \times (\overrightarrow{\theta_{nu}} \cdot \overrightarrow{x_{nu,rel}})\right)$$

Training: Similar loss function, Similar algorithm as LR

How it works?

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

How it works?

Satellite: 酒店位置好

Nuclei: 但是设施糟糕

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

How it works?

θ_{sat}: θ(酒店) θ(位置)

$x_{sat,rel}$: 酒店 位置   好

θ_{nu}: θ(但是) θ(糟糕)

$x_{nu,rel}$: 但是 设施 糟糕

Satellite: 酒店位置好

Nuclei: 但是设施糟糕

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

How it works?

Score(sen)

$\alpha_{sat,rel}$ $\alpha_{nu,rel}$

Score(sat) Score(nu)

θ_{sat}: θ(酒店) θ(位置)

θ_{nu}: θ(但是) θ(糟糕)

$x_{sat,rel}$: 酒店 位置 好

$x_{nu,rel}$: 但是 设施 糟糕

Satellite: 酒店位置好

Nuclei: 但是设施糟糕

Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

How it works?

Score(sen) $\xrightarrow{\text{sigmoid}}$ Category

$\alpha_{sat,rel}$ $\alpha_{nu,rel}$

Score(sat) Score(nu)

θ_{sat}: θ(酒店) θ(位置) θ_{nu}: θ(但是) θ(糟糕)

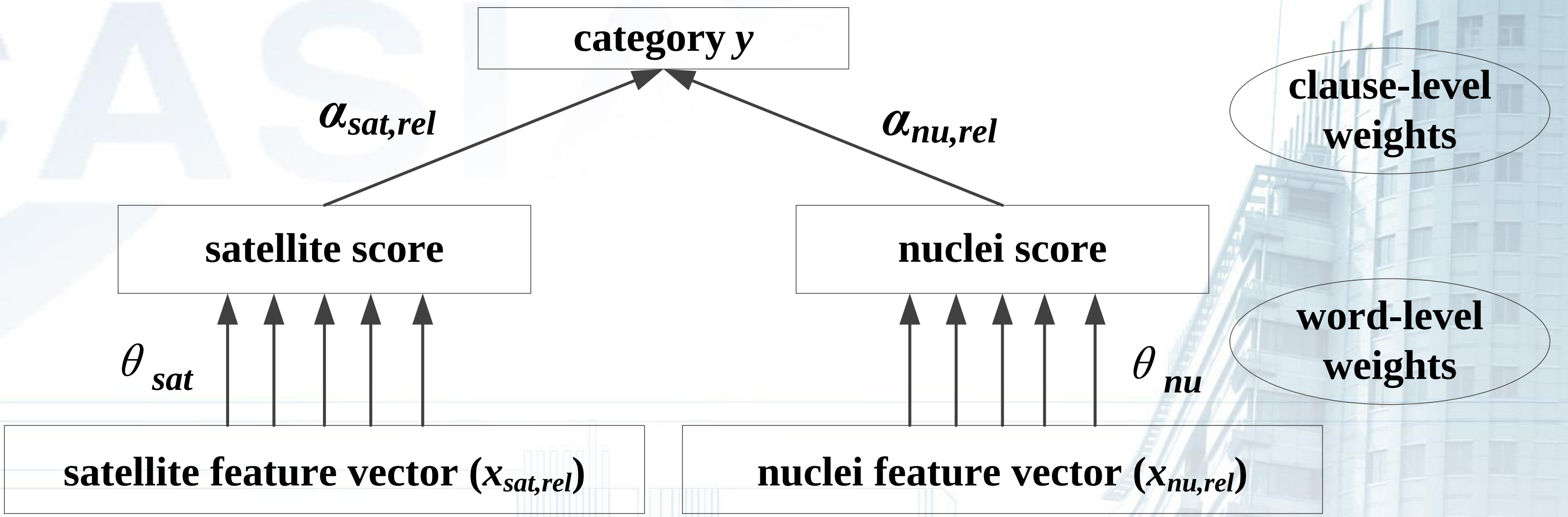
$x_{sat,rel}$: 酒店 位置 好 $x_{nu,rel}$: 但是 设施 糟糕

Satellite: 酒店位置好

Nuclei: 但是设施糟糕

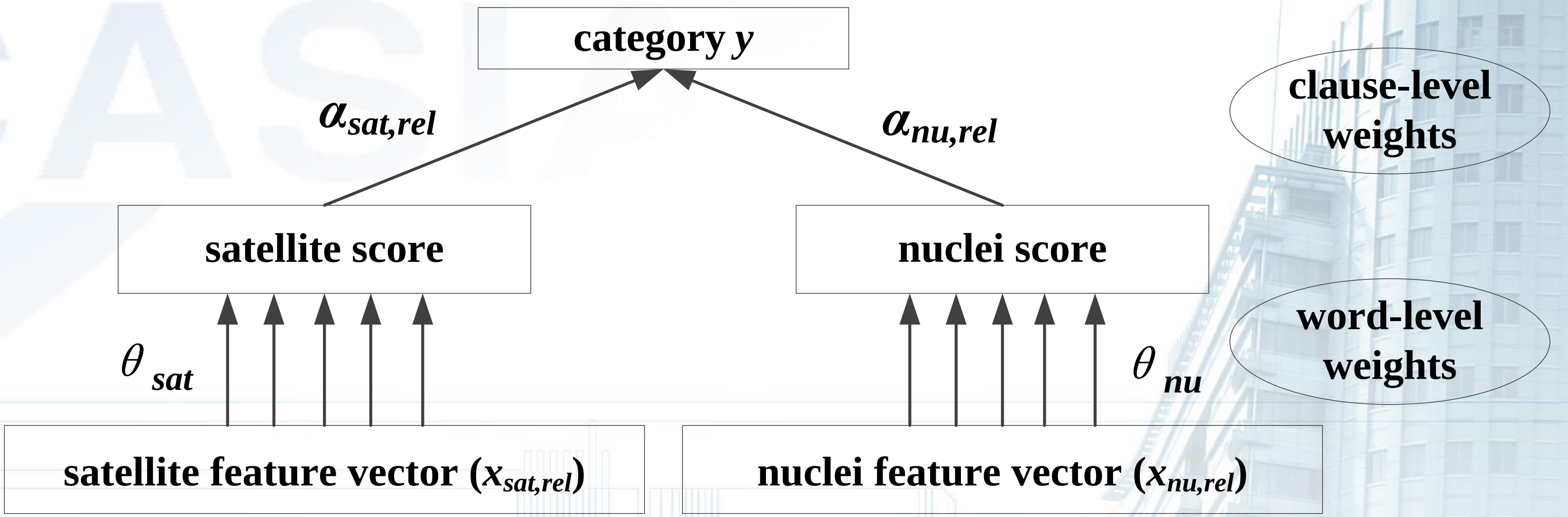
Input Sentence: 酒店位置好(The hotel location is good)，但是设施糟糕 (but its facilities are poor)

Parameters



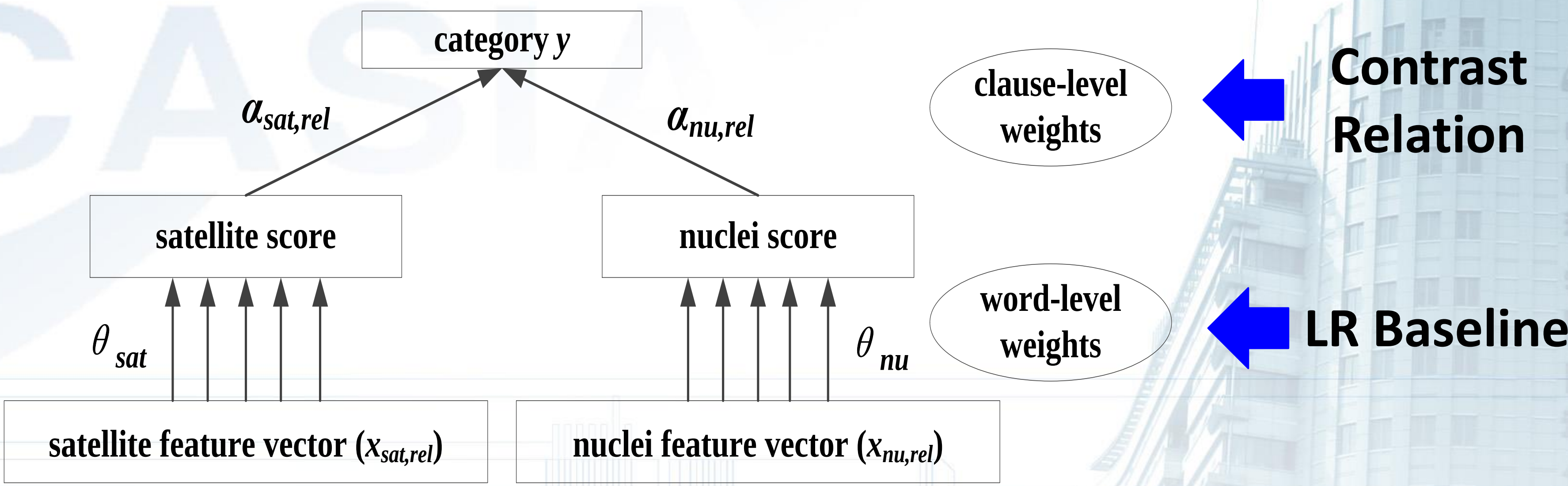
■ sat: satellite part, nu: nuclei part

Parameters



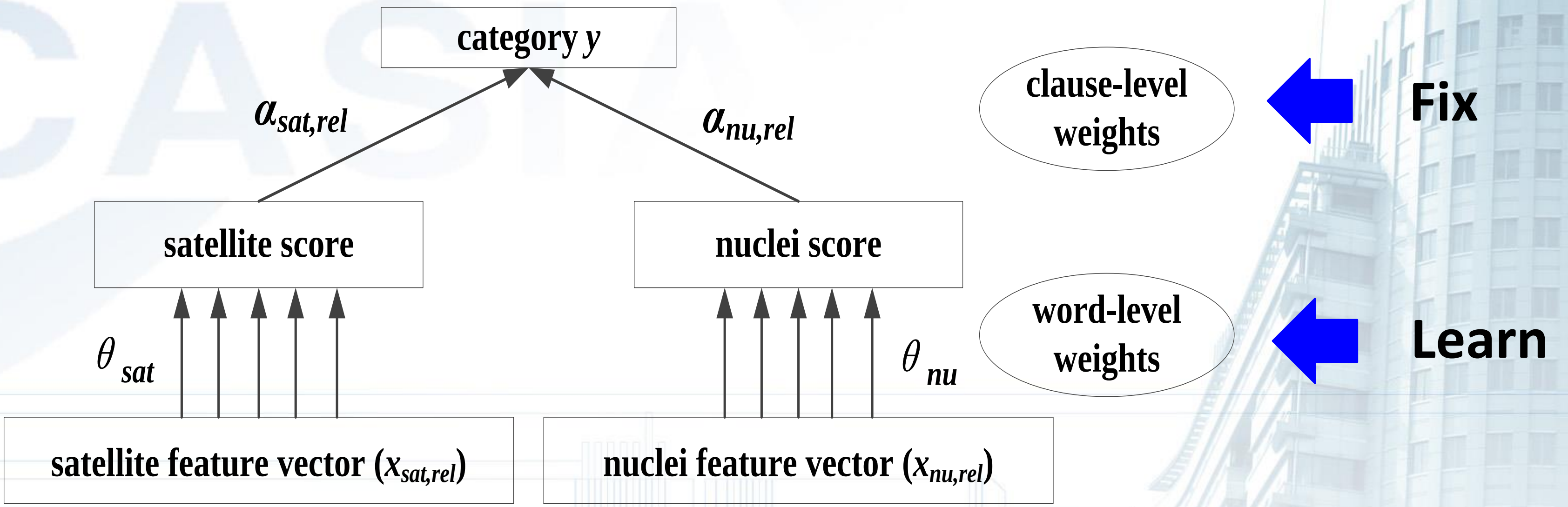
- sat: satellite part, nu: nuclei part
- rel: relation or connectives
  - Relation: 转折(contrast)
  - Connectives: 但是(but)，只是(yet)

Compared With Related Work



- SNSS: learn different clauses weight according to relation.

Compared With Related Work



- SNSS: learn different clauses weight according to relation.
- Joint strategy: simultaneity learns both two BOW model (satellite part and nuclei part).

Outline

- Motivation
- Related Work
- Methodology
- Approach Overview
- **Experiments**
- Conclusions

Distribution of Connectives

- Contrast sentences:
  - We get 693,125 sentences from Jing Dong.
  - Using contrast connectives, we get 41,629 contrast sentences, accounting for 6% in all sentences.

Experiment Dataset

- 2,000 contrast sentences as our experiment dataset:
  - Positive: 1,000
  - Negative: 1,000

Experiment Setting

- Model:
  - **Baseline**: SVM and LR

Experiment Setting

- Model:
  - **Baseline**: SVM and LR
  - **Other Models**:
    - SNSS (Single Nucleus Single Satellite Method), we get clause-level weights according to different clauses in contrast relation.
    - JS (joint strategy), the joint strategy used in [5].

Experiment Setting

- Model:
  - **Baseline**: SVM and LR
  - **Other Models**:
    - SNSS (Single Nucleus Single Satellite Method), we get clause-level weights according to different clauses in contrast relation.
    - JS (joint strategy), the joint strategy used in [5].
  - **Our Model**: we get the results of TLLR model.

Experiment Result

Systems	5-Fold Cross Validation Results (%)					
	1st	2nd	3rd	4th	5th	Average
SVM	86.00	86.00	83.00	85.00	85.50	85.10
LR	86.75	86.75	83.00	87.00	86.75	86.05
SNSS	86.25	88.00	83.50	86.00	86.25	86.00
JS	87.50	85.75	85.25	85.75	86.25	86.10
TLLR	86.58	87.70	84.95	86.73	87.73	86.74

Experiment Result

Systems	5-Fold Cross Validation Results (%)					
	1st	2nd	3rd	4th	5th	Average
SVM	86.00	86.00	83.00	85.00	85.50	85.10
LR	**86.75**	**86.75**	**83.00**	**87.00**	**86.75**	**86.05**
SNSS	86.25	88.00	83.50	86.00	86.25	86.00
JS	87.50	85.75	85.25	85.75	86.25	86.10
TLLR	86.58	87.70	84.95	86.73	87.73	86.74

- LR as our baseline system

Experiment Result

Systems	5-Fold Cross Validation Results (%)					
	1st	2nd	3rd	4th	5th	Average
SVM	86.00	86.00	83.00	85.00	85.50	85.10
LR	**86.75**	**86.75**	**83.00**	**87.00**	**86.75**	**86.05**
SNSS	**86.25**	**88.00**	**83.50**	**86.00**	**86.25**	**86.00**
JS	**87.50**	**85.75**	**85.25**	**85.75**	**86.25**	**86.10**
TLLR	86.58	87.70	84.95	86.73	87.73	86.74

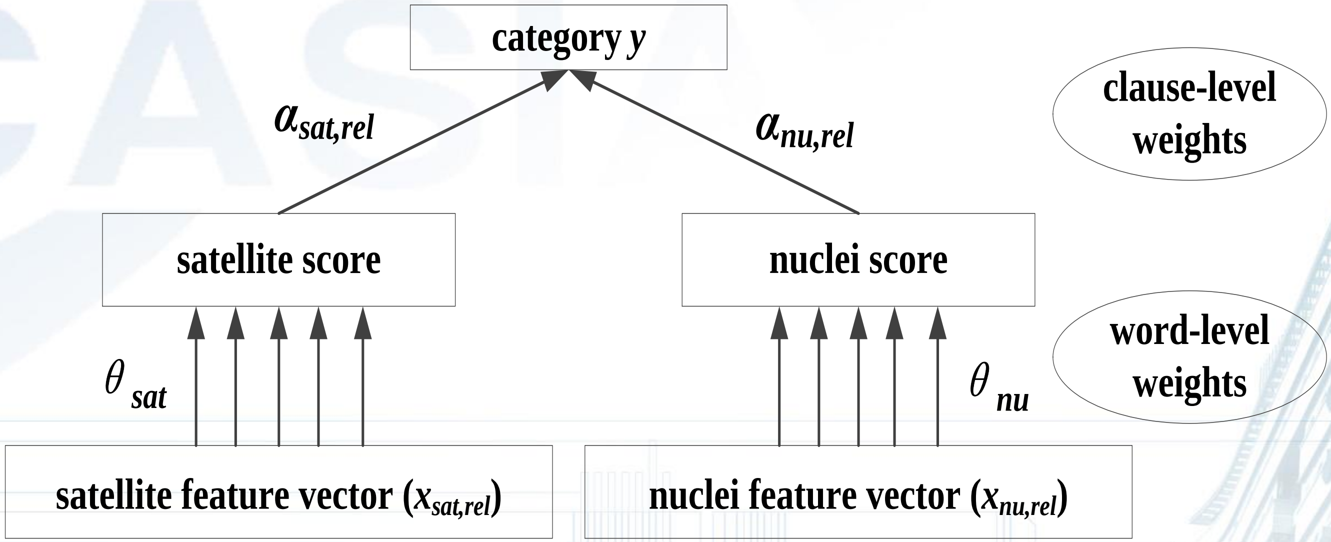
- LR as our baseline system
- Comparable Related Work: Word weight learn from sentence may not be suitable for clause.

Experiment Result

Systems	5-Fold Cross Validation Results (%)					
	1st	2nd	3rd	4th	5th	Average
SVM	86.00	86.00	83.00	85.00	85.50	85.10
LR	**86.75**	**86.75**	**83.00**	**87.00**	**86.75**	**86.05**
SNSS	86.25	88.00	83.50	86.00	86.25	86.00
JS	87.50	85.75	85.25	85.75	86.25	86.10
TLLR	**86.58**	**87.70**	**84.95**	**86.73**	**87.73**	**86.74**

- LR as our baseline system
- Comparable Related Work: Word weight learn from sentence may not be suitable for clause.
- Learn word weight and clause weight together is better

Parameter Analysis

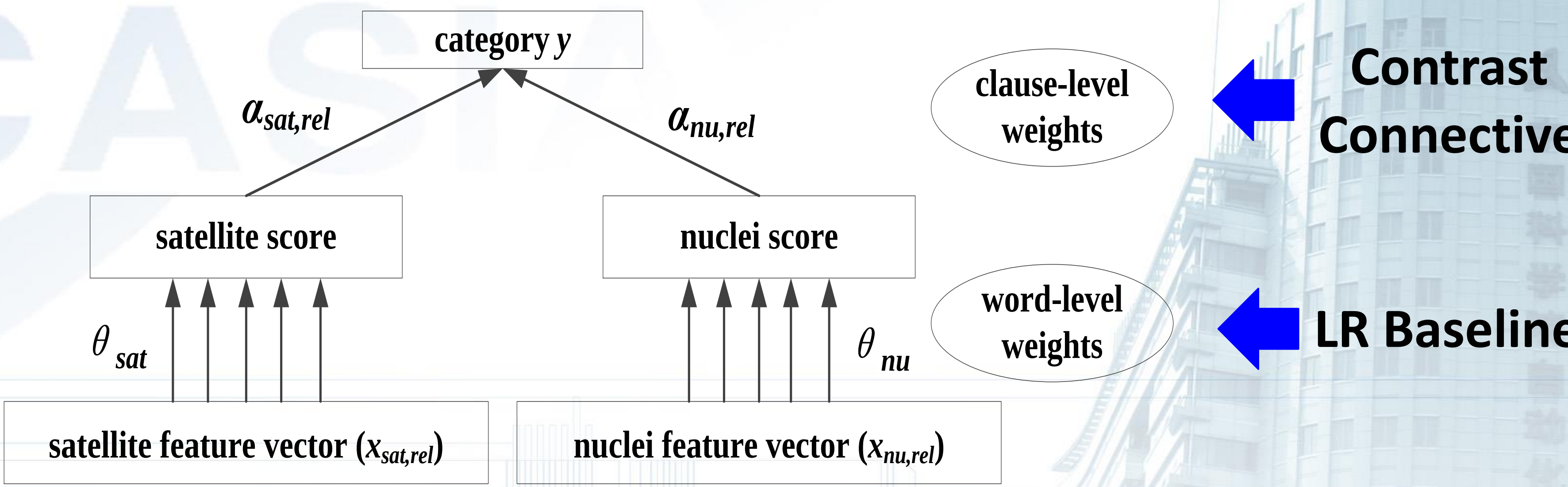


Contrast Relation

LR Baseline

Parameter 1: FixRelation

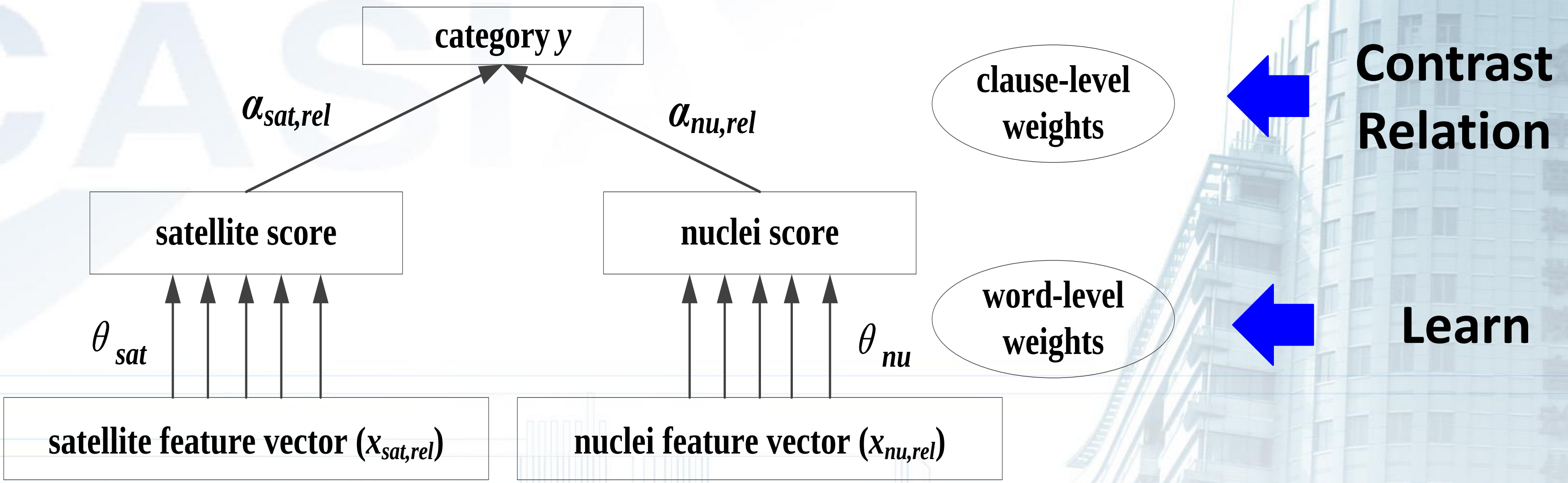
Parameter Analysis



Parameter 1: FixRelation

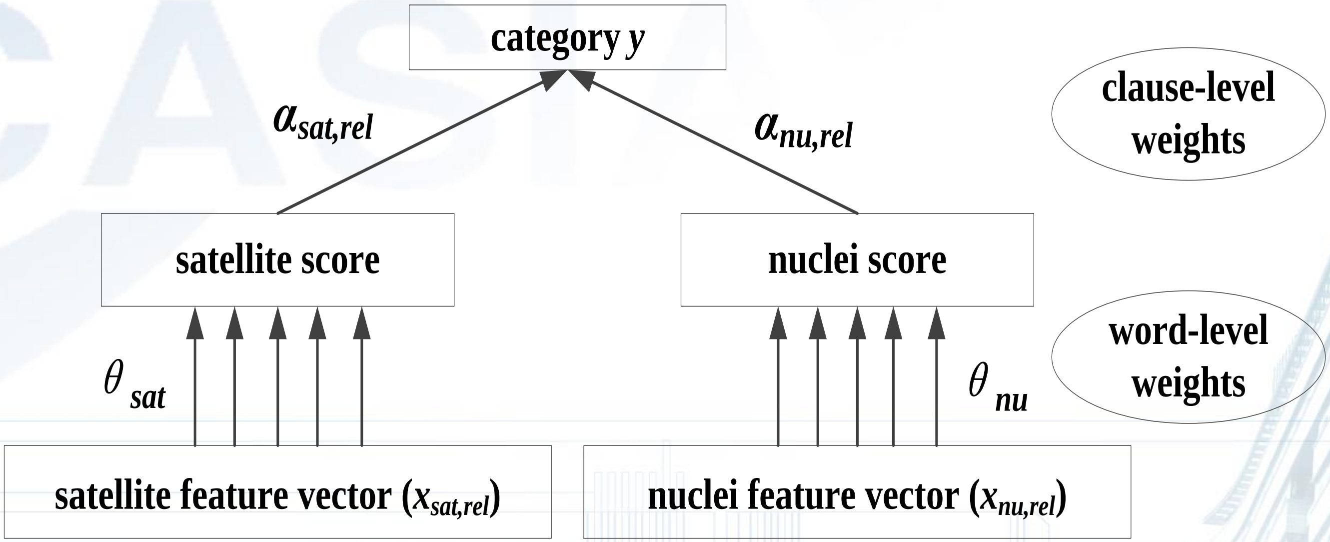
Parameter 2: FixConnective

Parameter Analysis



Parameter 1: FixRelation
Parameter 2: FixConnective
Parameter 3: DiffRelation

Parameter Analysis



Contrast Connective

Learn

Parameter 1: FixRelation
Parameter 2: FixConnective
Parameter 3: DiffRelation
Parameter 4: DiffConnective

Parameter Analysis Result

Systems	Average (5-fold)
FixRelation	86.00%
FixConnective	86.45%
DiffRelation	86.40%
DiffConnective	86.74%

- "Connective" is better than "Relation"

Parameter Analysis Result

Systems	Average (5-fold)
FixRelation	86.00%
FixConnective	86.45%
DiffRelation	86.40%
DiffConnective	86.74%

- "Connective" is better than "Relation"
- "Diff" is better than "Fix"

Parameter Analysis Result

Systems	Average (5-fold)
FixRelation	86.00%
FixConnective	86.45%
DiffRelation	86.40%
DiffConnective	**86.74%**

- "Connective" is better than "Relation"
- "Diff" is better than "Fix"
- "Diff" + "Connective" is the best

Clause-level Weight Analysis

Model	Type	Satellite	Nuclei
FixRelation	**转折|Contrast**	**0.48**	**0.52**
FixConnective	不过|however	0.6	0.4
	但|but	0.46	0.54
	但是|but	0.41	0.59
	只是|yet	0.56	0.44
	可是|however	0.41	0.59
DiffRelation	**转折|Contrast**	**0.45**	**0.55**
DiffConnective	不过|however	0.58	0.42
	但|but	0.42	0.58
	但是|but	0.36	0.64
	只是|yet	0.55	0.45
	可是|however	0.34	0.66

转折:

Nuclei > Satellite

Clause-level Weight Analysis

Model \ Type		Satellite	Nuclei
FixRelation	转折|Contrast	0.48	0.52
FixConnective	不过|however	0.6	0.4
	但|but	**0.46**	**0.54**
	但是|but	**0.41**	**0.59**
	只是|yet	0.56	0.44
	可是|however	**0.41**	**0.59**
DiffRelation	转折|Contrast	0.45	0.55
DiffConnective	不过|however	0.58	0.42
	但|but	**0.42**	**0.58**
	但是|but	**0.36**	**0.64**
	只是|yet	0.55	0.45
	可是|however	**0.34**	**0.66**

转折:

Nuclei > Satellite

但 但是 可是:

Nuclei > Satellite

Clause-level Weight Analysis

Model	Type	Satellite	Nuclei
FixRelation	转折|Contrast	0.48	0.52
FixConnective	**不过|however**	**0.6**	**0.4**
	但|but	0.46	0.54
	但是|but	0.41	0.59
	只是|yet	**0.56**	**0.44**
	可是|however	0.41	0.59
DiffRelation	转折|Contrast	0.45	0.55
DiffConnective	**不过|however**	**0.58**	**0.42**
	但|but	0.42	0.58
	但是|but	0.36	0.64
	只是|yet	**0.55**	**0.45**
	可是|however	0.34	0.66

转折:

Nuclei > Satellite

但 但是 可是:

Nuclei > Satellite

不过 只是:

Nuclei < Satellite

Outline

- Motivation
- Related Work
- Approach Overview
- Methodology
- Experiments
- **Conclusions**

Conclusions

- TLLR model is proposed.
  - TLLR model considers clause-level weights with connectives, not just relations.
  - TLLR model can learn word weights and clause-level weights together.
- Future Work:
  - Consider other relations.
  - Deduce the overall polarity from the polarity of each aspect.

References

- [2] Bas Heerschop, Frank Goossen, Alexander Hogenboom, Flavius Frasincar, Uzay Kaymak, and Franciska de Jong.2011. Polarity analysis of texts using discourse structure. *In Proceedings of the ACM international conference on Information and knowledge management(CIKM). pages1061–1070.*
- [5] Shoushan Li and Chu-Ren Huang. 2009. Sentiment Classification Considering Negation and Contrast Transition.*In Proceedings of the Pacific Asia Conference on Language, Information, and Computation (PACLIC).*
- [11] Maite Taboada, Kimberly Voll, and Julian Brooke. 2008. Extracting sentiment as a function of discourse structure and topicality. Simon Fraser University, Tech. Rep, 20.
- [13] Fei Wang, Yunfang Wu, and Likun Qiu. 2012. Exploiting Discourse Relations for Sentiment Analysis. *In Proceeding of the Intenerational Conference of Computational Linguistics(COLING). pages 1311–1319.*

Thanks!

Distribution of Connectives

- Contrast sentences:
  - We get 693,125 sentences from Jing Dong.
  - Using contrast connectives, we get 41,629 contrast sentences, accounting for 6% in all sentences.

Contrast Connectives	Sentence Number	Sentence Proportion (%)
不过(however)	14,869	35.72
但是(but)	11,919	28.63
但(but)	9,553	22.95
只是(yet)	3,047	7.32
可是(however)	2,241	5.38
All	41,629	100