Opinion Extraction

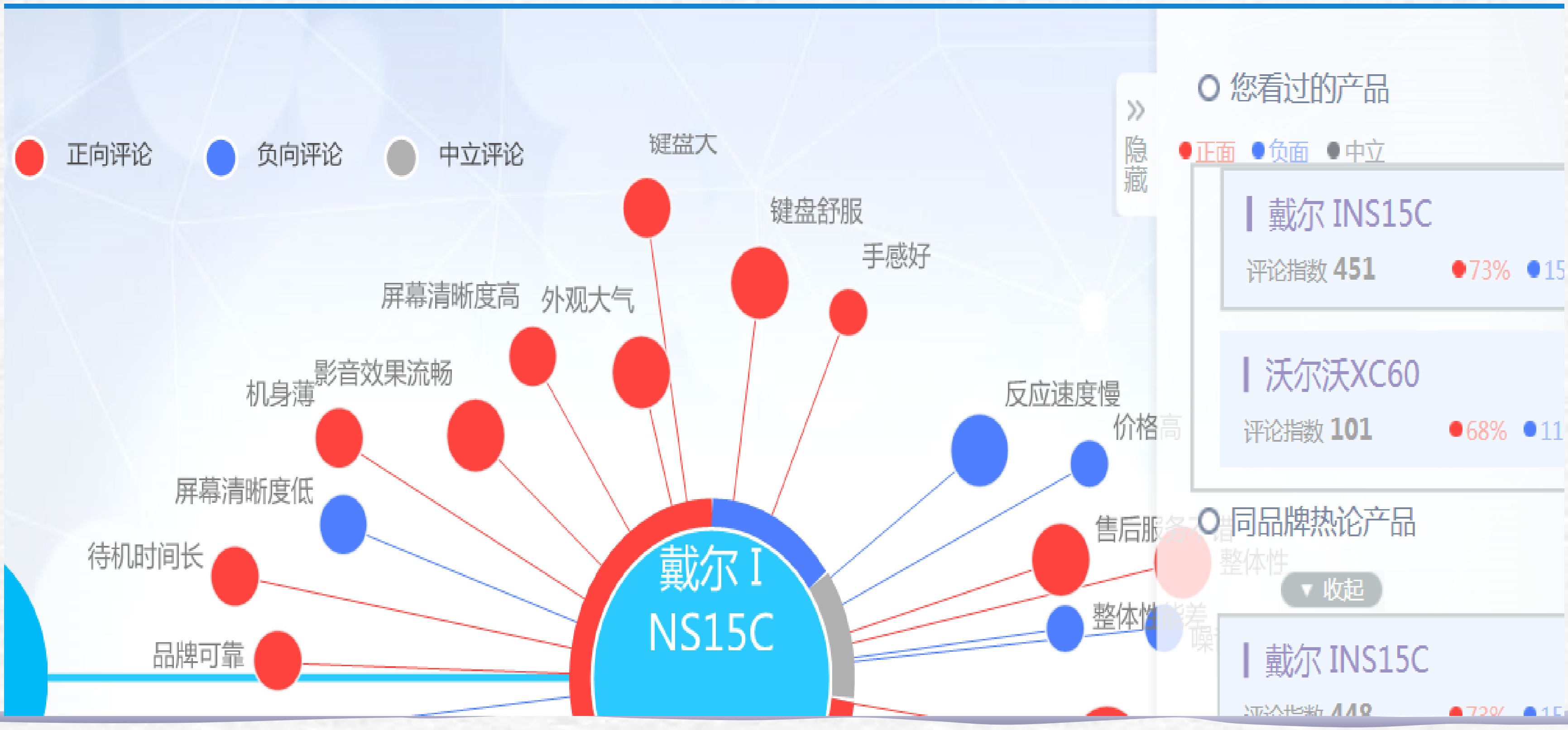
Fang Li

Contents

- Research Background
- Opinion Definition
- Opinion Representation
- Opinion Extraction
- Recent Developments

Research Background

- Sentiments about a product (产品的情感分析)
- Extraction of attributes of a product (产品的属性抽取)



Research Background (from Bing Liu)

- "Opinions" are key influencers of behaviors.
- Our perceptions of reality and decisions of action are largely conditioned on how others see the world or their experiences.

→ opinion mining

Research Background

- Movie: is this review positive or negative?
- Products: what do people think about the new iPhone?
- Public sentiment: how is consumer confidence? Is despair increasing?
- Politics: what do people think about this candidate or issue?
- Prediction: predict election outcomes or market trends from sentiment

What is an opinion?

1) A view or judgment formed about something, not necessarily based on fact or knowledge.

2) According to (Wiebe et al., 2005):

"Subjective expressions are words and phrases being used to express opinions, emotions, evaluations, speculations, etc.

Research Background

Two kinds of information on the Web:

- **Factual Information**: entity, relation, event.
- **Subjectivity information**: opinions.

"I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, …"

*The 1st sentence is **a fact**, the 2, 3,4 and 5 are **opinions**. They are positive for the iphone.*

Explicit and Implicit Opinion

- An explicit opinion on feature f is an opinion explicitly expressed on f in a subjective sentence.

"The voice quality of this phone is amazing."

- An implicit opinion on feature f is an opinion on f implied in an objective sentence.

"The earphone broke in two days."

Explicit and Implicit Aspects

- Explicit aspects: Aspects explicitly mentioned as nouns or noun phrases in a sentence.

"*The picture quality is of this phone is great.*"

- Implicit aspects: Aspects not explicitly mentioned in a sentence but are implied.

"This car is so expensive." price

"This phone will not easily fit in a pocket. size

Two types of opinions

- Regular opinion

"the **touch screen** is really cool"

"after taking the drug, my pain has gone"

implicit feature

Feature: the effect of the drug

Sentiment: positive.

"The earphone broke in two days."

implicit opinion

- Comparative opinion

"It is much better than my old Blackberry"

expresses a relation of similarities or differences between two or more objects(features)

Model of an Opinion: / Opinion Representation

An opinion is a quintuple (Oj, Fjk, Sijkl, Hi,Tl):
Where

- O_j: is an object, a target entity.
- F_{jk} : is a feature of an object O_j, or an aspect of a target entity O_j
- S_{ijkl}: the sentiment value of the opinion from opinion holder H_i on feature F_{jk} of object O_j at Time T_l . The value of S is neg.(-), pos.(+), neu, or more granular ratings.
- H_i: is an opinion holder
- T_l : is the time when opinion expressed.

Example

"I bought an iPhone a few days ago. It is such a nice phone. The touch screen is really cool. The voice quality is clear too. It is much better than my old Blackberry, which was a terrible phone and so difficult to type with its tiny keys. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive, …"

- (iphone, ? , nice +, I, t)
- (iphone, touch screen, cool+, I, t)
- (iphone, voice quality, clear+, I, t)
- (iphone, price!, expensive -, my mother, t)

Implicit feature

Tasks of opinion extraction

- Simplest task:

Is the attitude of this text positive or negative? → sentiment classification

- More complex:

Rank the attitude of this text from 1 to 5

→ ranking problem

- Advanced:

Detect the target, source, or complex attitude types → information extraction

Sentiment Classification

- Document-level sentiment classification
- Sentence level sentiment classification

Document-Level Sentiment Classification

- **Assumption:** d expresses opinions on single object o and the opinions are from a single holder h.
- **Task**: given a quintuple (o, f, so, h, t), where f = o and h, t are assumed to be known or irrelevant, find the value of so.
- **Methods**:

supervised learning & unsupervised learning.

Classification based on supervised learning

Models:

- Naïve Bayesian, Support Vector Machines

Features:

- Terms and their frequency: n-gram and their counts.
- Part of speech tags: adjectives are important indicators for subjectivities and opinions.
- Opinion words and phrases: wonderful, poor, hate, like,…
- Syntactic dependency: features generated from parsing or dependency trees are also used.
- Negation: change the opinion orientation.

Classification based on unsupervised learning

- Step1: Extracts phrases containing adjectives or adverbs.
- Step2: Estimates the orientation of the extracted phrases using PMI （Pointwise Mutual Information）:

$$PMI(term_1, term_2) = \log_2\left(\frac{\Pr(term_1 \wedge term_2)}{\Pr(term_1)\Pr(term_2)}\right).$$

$$oo(phrase) = PMI(phrase, \text{“excellent”}) - PMI(phrase, \text{“poor”}).$$

- Step3: Computes the average oo/so of all phrases in the review, and classifies the review as positive or negative.

About PMI (pointwise mutual information 点互信息)

- Evaluate the relevance between two variables. 用来衡量两个事件,变量,**两个词**之间的**相关性**.

$$\mathrm{PMI}(x;y) = \log\frac{p(x,y)}{p(x)p(y)} = \log\frac{p(x|y)}{p(x)} = \log\frac{p(y|x)}{p(y)}$$

- If x and y relevant, the joint probability p(x,y) will be much larger than chance P(x)P(y) and PM I(x,y) >>0.
- if x and y not relevant, then PMI(x,y) =0
- if x and y are in complementary distribution, then PMI(x,y) <<0

Example for PMI

		Count(w, context)			
	computer	data	pinch	result	sugar
apricot	0	0	1	0	1
pineapple	0	0	1	0	1
digital	2	1	0	1	0
information	1	6	0	4	1

p(w=information, c=data) = 6/20=.3

p(w=information) = 12/20=.6

p(c=data) = 7/20=.35

P(w=information,c=sugar)=1/20=0.05

P(c=sugar)=3/20=0.15

PMI (information,data) = log 0.3/ (0.6*0.35)=log 1.43=0.15

PMI(information,sugar)= log 0.05/(0.6*0.15)=log 0.55= -0.26

Document-Level Sentiment Classification

Advantage:

- Coarse-grained analysis
- Detection of a general sentiment trend of a document

Problems:

- Different polarities towards different features, e.g. *This film should be brilliant. The characters are appealing. Stallone plays a happy, wonderful man. His sweet wife is beautiful and adores him. He has a fascinating gift for living life fully. It sounds like a great story, **however, the film is a failure**.*

Sentence-Level Classification

Task: Given a sentence s, two sub-tasks are performed:

1. Subjectivity Classification: Determine whether s is a subjective sentence or an objective sentence.

2. Sentiment Classification: If s is a subjective sentence, determine whether it expresses a positive or negative opinion.

Sentence-Level Classification

Advantage:

- More specific than document-level analysis
- The results can be reused as input for document-level classification

Problems:

- Multiple sentiment expressions with different polarities, e.g.*The very brilliant organizer failed to solve the problem.*

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. **Thumbs up? Sentiment Classification using Machine Learning Techniques.** EMNLP-2002, 79—86.
Bo Pang and Lillian Lee. 2004. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. ACL, 271-278
Bo Pang and Lillian Lee, Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales, Proceedings of ACL 2005

Sentiment Classification in Movie Reviews
（a baseline algorithm）

- Polarity detection:
  - Is an IMDB movie review positive or negative?
- Data: *Polarity Data 2.0:*
  - http://www.cs.cornell.edu/people/pabo/movie-review-data

Sentiment Classification in Movie Reviews

Fig.1: human-based classifier
Fig.2: a list of seven positive and seven negative words

percentage of documents where the two sentiments were rated equally likely

it is worthwhile to explore corpus-based techniques, rather than relying on prior intuitions, to select good indicator features and to perform sentiment classification in general.

	Proposed word lists	Accuracy	Ties
Human 1	positive: *dazzling, brilliant, phen* negative: *suck, terrible, awful, unw*		75%
Human 2	positive: *gripping, mesmerizing, ri* *awesome, thrilling, badass, ex* negative: *bad, cliched, sucks, borin*		39%

Figure 1: Baseline results for human word lists. Data: 700 positive and 700 negative reviews.

	Proposed word lists	Accuracy	Ties
Human 3 + stats	positive: *love, wonderful, best, great, superb, still, beautiful* negative: *bad, worst, stupid, waste, boring, ?, !*	69%	16%

ure 2: Results for baseline using introspection and simple statistics of the data (including *test* da

What are the conclusions based on the experiments?

	Features	# of features	frequency or presence?	NB	ME	
(1)	unigrams	16165	freq.	**78.7**	N/A	72.8
(2)	unigrams	”	pres.	81.0	80.4	**82.9**
(3)	unigrams+bigrams	32330	pres.	80.6	80.8	**82.7**
(4)	bigrams	16165	pres.	77.3	**77.4**	77.1
(5)	unigrams+POS	16695	pres.	81.5	80.4	**81.9**
(6)	adjectives	2633	pres.	77.0	**77.7**	75.1
(7)	top 2633 unigrams	2633	pres.	80.3	81.0	**81.4**
(8)	unigrams+position	22430	pres.	81.0	80.1	**81.6**

Figure 3: Average three-fold cross-validation accuracies, in percent. Boldface: best performance for a given setting (row). Recall that our baseline results ranged from 50% to 69%.

Baseline Conclusion

- Unigram as features are better, especially of top unigrams.
- Feature frequency are not better than presence.
- Part of speech does not help a lot in sentiment analysis.
- Adjectives does not help a lot.
- SVM is better than other classifiers.

Advanced Task

- Identify the targets of opinions, i.e. Entities and their aspects, or object and their features.
- Reasons:
- Without knowing targets, opinions are of limit use.
- Target-opinion pair: more precise, more useful.
- Implicit, missing features or targets: more difficulty task.

Opinion Extraction

An opinion is represented as a quintuple (O_j, F_{jk}, S_{ijkl}, H_i,T_l):
Where:

- O_j: is an object, a target entity. named entity extraction +
- F_{jk} : is a feature of an object O_j, or an aspect of a target entity O_j information extraction
- S_{ijkl}: the sentiment value of the opinion from opinion holder H_i on feature F_{jk} of object O_j at Time T_l . The value of S is neg.(-), pos.(+), neu, or more granular ratings.
- H_i: is an opinion holder information/data extraction
- T_l : is the time when opinion expressed. information/data extraction

✓Co-reference Resolution

✓Synonym Match (voice = sound quality)

Steps of Opinion Mining

1. Source the data, e.g., reviews, blogs, etc
   - (1) Crawl all data, store and search them, or (2) Crawl only the target data
2. Extract the right entities & aspects
   - Group entity and aspect expressions,

     Moto = Motorola, photo = picture, etc …
3. Aspect(feature)-based opinion mining

   Discover all quintuples (Store the quintuples in a database)

Aspect-based Opinion Mining

- Object-specific: reviews for some known entities or events. Reviewers simply express positive and negative opinions on different aspects of the entity.
- Object unknown: for blogs, forum discussions, both entity and aspects are unknown.
- <u>Four methods</u> will be discussed.

Aspect extraction (1)

Hu and Liu, 2004: M. Hu and B. Liu, "Mining and summarizing customer reviews," Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), pp. 168–177, 2004.

- A frequency-based approach (Hu and Liu, 2004): nouns (NN) that are frequently talked about are likely to be true aspects (called frequent aspects) .
- Why the frequency based approach?
- Different reviewers tell different stories.
- When product aspects/features are discussed, the words they use converge.
- They are the main aspects.
- How to find frequent nouns and noun phrases?

POS tagger + TF(词频)-IDF(逆文档频率)

Aspect Extraction(1): Two-steps

- Step 1: Finding <u>frequent</u> nouns and noun phrases. → many noises

- Step2: Finding infrequent features by making use of opinion words:

"*The **<u>pictures</u>** are absolutely amazing.*"

"*The **software** is amazing.*"

Aspect Extraction(1): cont.

- Using PMI measure to remove those noun phrases that may not aspects.

$$PMI(f,d) = \frac{hits(f \wedge d)}{hits(f)hits(d)}$$

- f is a candidate noun phrase identified in step 1 and d is a discriminator. E.g. a scanner class, d: "of scanner", "scanner has", "scanner comes with",

Aspect Extraction (2)

- Key idea: opinions have targets, i.e., opinion words are used to modify aspects and entities.

“The pictures are absolutely amazing.”

“This is an amazing software.”

- <u>the nearest noun</u> to the opinion word.

Aspect Extraction (3)
Main idea of the method

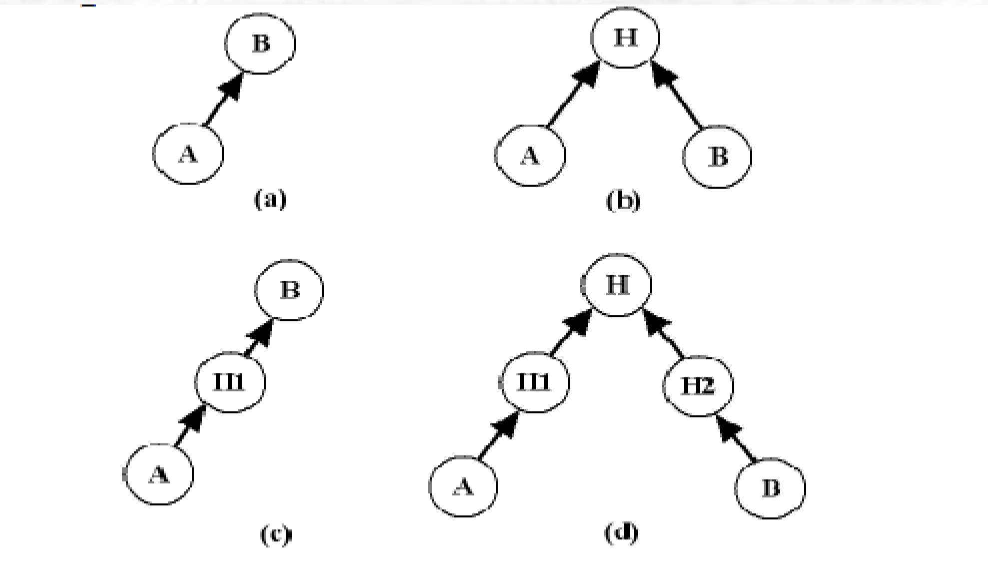
G. Qiu, B. Liu, J. Bu and C. Chen. Expanding Domain Sentiment Lexicon through Double Propagation, International Joint Conference on Artificial Intelligence (IJCAI-09), 2009.

- Exploit certain syntactic relations of opinion words and object features for extraction.
- **Opinion words** can be recognized by identified **features**, and **features** can be identified by **known opinion words.**
- The extracted opinion words and features are utilized to identify new opinion words and new features, which are used again to extract more opinion words and features.

Aspect Extraction (3): cont.

Relations of Sentiment Words and Features

Both A and B can be sentimental words or features



(a)(b): A *direct relation* means that one word depends on the other word directly or they both depend on a third word directly.

(c)(d): An *indirect relation* means that one word depends on the other word through other words or they both depend on a third word indirectly.

Aspect Extraction (3) cont.

O,T are known opinion words and targets.
o,t are unknown opinion words and targets.

4 tasks:

- R1: to extract target (t) using opinion words (O)
- R2: to extract opinion words (o) using targets (T)
- R3: to extract targets (t) using extracted targets (T).
- R4: to extract opinion words (o) using known opinion words (O)

Aspect Extraction (3) cont.

RuleID	Observations	output	Examples
$R1_1$	$O \rightarrow O\text{-}Dep \rightarrow T$ s.t. $O \in \{O\}$, $O\text{-}Dep \in \{MR\}$, $POS(T) \in \{NN\}$	$t = T$	The phone has a good "screen". ($good \rightarrow mod \rightarrow screen$)
$R1_2$	$O \rightarrow O\text{-}Dep \rightarrow H \leftarrow T\text{-}Dep \leftarrow T$ s.t. $O \in \{O\}$, $O/T\text{-}Dep \in \{MR\}$, $POS(T) \in \{NN\}$	$t = T$	"iPod" is the best mp3 player. ($best \rightarrow mod \rightarrow player \leftarrow subj \leftarrow iPod$)
$R2_1$	$O \rightarrow O\text{-}Dep \rightarrow T$ s.t. $T \in \{T\}$, $O\text{-}Dep \in \{MR\}$, $POS(O) \in \{JJ\}$	$o = O$	same as $R1_1$ with screen as the known word and good as the extracted word
$R2_2$	$O \rightarrow O\text{-}Dep \rightarrow H \leftarrow T\text{-}Dep \leftarrow T$ s.t. $T \in$ $\{T\}$, $O/T\text{-}Dep \in \{MR\}$, $POS(O) \in \{JJ\}$		same as $R1_2$ with iPod as the known word and best as the extract word
$R3_1$	$T_{i(j)} \rightarrow T_{i(j)}\text{-}Dep$ … $Dep \in \{C$ …		Does the player play dvd with … and "video"? ($video \rightarrow$ … dio)
$R3_2$	T …		… s a great lens. … $s \leftarrow subj \leftarrow G3$)
$R4_1$			… zing and … $\rightarrow conj \rightarrow$
$R4_2$	$\{C$ …		… uy a sexy, "cool", … ilable mp3 player, … oose iPod. ($sexy \rightarrow$ $player \leftarrow mod \leftarrow cool$)

1) If "observations" then output.
2) Underlined word is the known word,
3) Double quoted word is the target outputted.
4){MR}:dependency relations includes (mod,pnmod,subj,sobj,obj2,..)

Example

- Given a text *"Canon G3 takes <u>great</u> pictures, The picture is amazing, You may have to get more storage to store high quality pictures and recorded movies, and The software is amazing."*

Based on ($R1_1$): great→"picture"

($R2_2$): picture→"amazing"

($R3_1$): Picture and …"Movies" is a feature

($R1_2$): Amazing→ "software"

Aspect Mining (4)

- A mutual reinforcement method (Su et al. 2009)
- It proposed an unsupervised approach which exploits the mutual reinforcement rel… nship between aspects and opinion w…
- it uses th… …inion word…
- The … aspects a… …tely, but before clusterin… …lustering results of the other set is used to update the pairwise weight of the set.

Q. Su, X. Xu, et al. Hidden Sentiment Association in Chinese Web Opinion Mining. Proceedings of WWW’08, pp. 959-968, 2008.

Aspect Mining (4)

1. MiniCooper Conve
MiniCooper Conver
2. 车身线条流畅，雍容
The car has a smoo
3. 初见辉腾给人的感觉
The first feeling wh
Germany car, ordin
4. 雪铁龙C5的前脸设计
Citrön C5's front
5. 05款的东方之子表现
Eastar 05 has good

6. *The Corvette C6 i*
7. *The NSX is truly*
8. *Parts are expensiv*
9. *It is superbly sport*
10. *Transmission is*

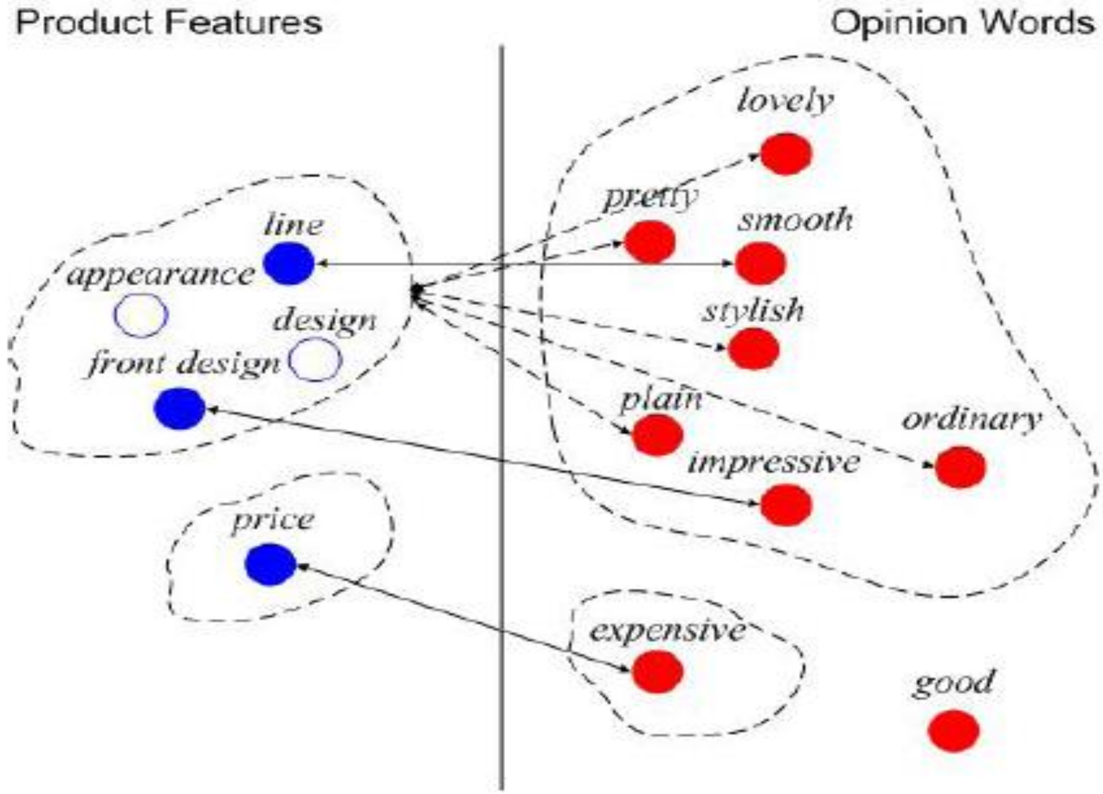


Figure 1: Complicated Relationship Between Product Features and Opinion Words in Real Reviews (solid circle/solid line represents an explicit word/relationship; hollow circle/dash line represents an implicit word/relationship)

After Aspects Extraction

- **Group Synonyms**: different words to describe an aspect. E.g. picture=photo for digital camera; picture=movie for movie review.

e.g: using an ontology:

- Camera
  - Lens
    - Digital Zoom
    - Optical Zoom
    - …
  - Editing/Viewing
    - Viewfi nder
    - …
  - Flash
- Image
  - Image Type
    - TIFF
    - JPEG
    - …
  - Resolution
    - Effective Pixels
    - Aspect Ratio
  - …

- **Mapping to implicit aspects**: e.g. heavy is general for weight, however, “the traffic is very heavy” is not.

Opinion Extraction and polarity identification

Lexicon-based approach:

- Positive words: better, good, super,…
- Negative words: bad, …
- Neutral words (context dependent): long,…

"the battery life is *long*, but the time taken to focus is *long*"

Polarity Identification

- Positive words +1, negative words-1, neutral 0
- Handling negations: revise the opinion scores
- But-clauses: the opinion orientation before but and after but are opposite to each other.
- Aggregating opinions: $score(f_i, s) = \sum_{op_j \in s} \frac{op_j.so}{d(op_j, f_i)}$,

where op_j is an opinion word in s, $d(op_j, fi)$ is the distance between feature fi and opinion word opj in s. opj.so is the orientation or the opinion score of opi.

Example

“The picture quality of this camera is not great(+1) , but the battery life is long(0).”

→ “The picture quality of this camera is not great(-1) , but the battery life is long(+1).”

Some rules:

Negation Neg → Positive

Negation Pos → Negative

Desired value range → Positive

Decreased Neg → Positive….

Comparative opinion

Three cases:

- "optics of camera A is better than that of camera B" --not equal
- "camera A and camera B both come in 7MP" --equal
- "camera A is the cheapest in market" --superlative

Task of Comparative opinion

- Input: Given an opinionated document d,
- Output: comparative opinions:

  (E1, E2, A, po, h, t),

E1 and E2: the entity sets being compared.

A: their shared aspects

Po: is the preferred entity set

H: the opinion holder

t is the time

Note: not positive or negative opinions.

Method of Comparative opinion mining

1. Identify comparative sentences

- Strong patterns involving comparative keywords: *Whereas/IN, but/CC, however/RB, while/IN, though/IN, etc*
- Supervised learning: classify into three types.

2. Extraction of different items

Rule based: objects or features are nouns and pronouns.

Machine learning: HMM, CRF

3. Determine preferred entities (opinions)

Parsing and opinion lexicon

Opinions Output

- A collection of opinions are valuable.
- Features-based summary are useful.

Cellular phone 1:

PHONE:

Positive: 125 <individual review sentences>

Negative: 7 <individual review sentences>

Feature: **voice quality**

Positive: 120 <individual review sentences>

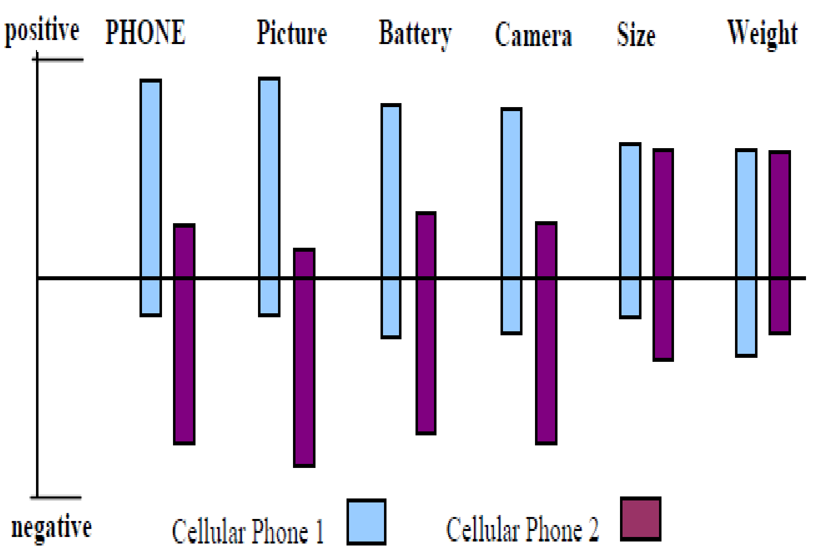
Negative: 8 <individual review sentences>

Feature: **size**

Positive: 80 <individual review sentences>

Negative: 12 <individual review sentences>

…



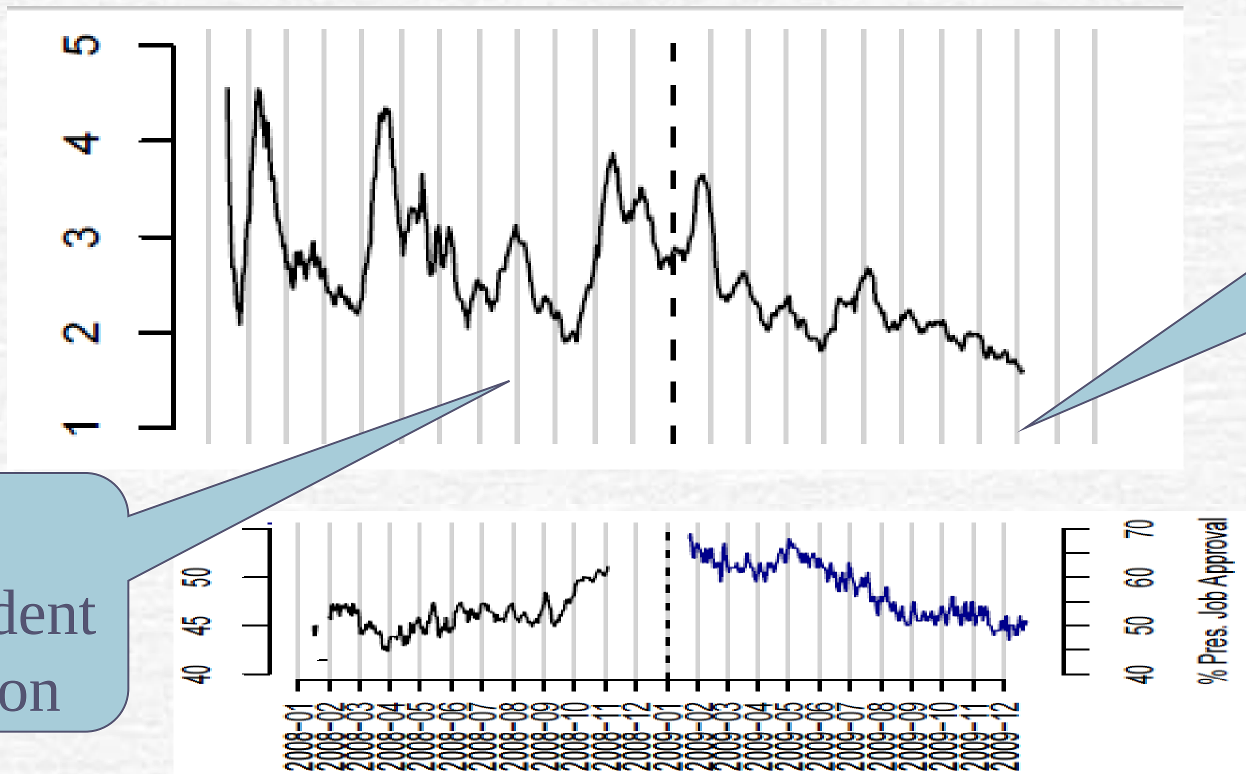
Opinions Output (cont.)

Structured opinion summary:

- **Feature buzz summary**: shows the relative frequency of feature mentions→ a company knows what the customer care about.
- **Object buzz summary**: shows the frequency of mentions of different competing products → customer knows which product is better.
- **Trend tracking**: add time dimension, get trend reports.

Trend Tracking Example

- Top: Public opinion measured from polls
- Bottom: Sentiment measured from text



Some Challenges

- Sarcastic sentences

e.g. “What a great car, it stopped working in the second day.” ×

- implicit opinion

e.g. Stephanie McMahon is the next Stalin √

- Opinion spamming

refers to people giving **fake or untruthful** opinions.

- Utility of opinions

refers to the usefulness or **quality** of opinions.

Summarization

- Opinion Definition and Representation
- Sentiment and subjectivity classification
- Aspect-based opinion mining
- Sentiment analysis of comparative sentences

Reference

- Bing Liu. Sentiment Analysis and Subjectivity. in Handbook of Natural Language Processing, Second Edition, (editors: N. Indurkhya and F. J. Damerau), 2010
- **A comprehensive bibliography of sentiment analysis:** http://www.cs.pitt.edu/~wiebe/subjectivityBib.html

Competitions about Sentiment Analysis

- Sentiment Analysis in Twitter (SemEval-2016)
- 43 teams joined the competition.
- Positive, Negative in general or about a given entity.

Two variants in 2016

- **Ordinal Classification**:

{HIGHLYPOSITIVE, POSITIVE, NEUTRAL, NEGATIVE, HIGHLYNEGATIVE

- **Quantification**: aggregate data

Supervised class prevalence estimation:

Estimating the distribution of the classes in a set of unlabelled items.

Tasks

1. **Subtask A:** Given a tweet, predict whether it is of positive, negative, or neutral sentiment.
2. **Subtask B:** Given a tweet known to be about a given topic, predict whether it conveys a positive or a negative sentiment towards the topic.
3. **Subtask C:** Given a tweet known to be about a given topic, estimate the sentiment it conveys towards the topic on a five-point scale ranging from HIGHLYNEGATIVE to HIGHLYPOSITIVE.
4. **Subtask D:** Given a set of tweets known to be about a given topic, estimate the distribution of the tweets in the POSITIVE and NEGATIVE classes.
5. **Subtask E:** Given a set of tweets known to be about a given topic, estimate the distribution of the tweets across the five classes of a five-point scale, ranging from HIGHLYNEGATIVE to HIGHLYPOSITIVE.

Evaluation Metrics for Task 1

$$F_1^{PN} = \frac{F_1^P + F_1^N}{2} \quad (1)$$

F_1^P is the F_1 score for the POSITIVE class:

$$F_1^P = \frac{2\pi^P \rho^P}{\pi^P + \rho^P} \quad (2)$$

Here, π^P and ρ^P denote precision and recall for the POSITIVE class, respectively:

$$\pi^P = \frac{PP}{PP + PU + PN}$$

$$\rho^P = \frac{PP}{PP + UP + NP}$$

		Gold Standard		
		POSITIVE	NEUTRAL	NEGATIVE
Predicted	POSITIVE	PP	PU	PN
	NEUTRAL	UP	UU	UN
	NEGATIVE	NP	NU	NN

Table 7: The confusion matrix for Subtask A. Cell XY stands for "the number of tweets that the classifier labeled X and the gold standard labells as Y". P, U, N stand for POSITIVE, NEUTRAL, NEGATIVE, respectively.

Results for Task 1

#	System	F_1^{PN}	ρ^{PN}	Acc
1	SwissCheese	$\mathbf{0.633}_{1}$	0.667_{2}	0.646_{1}
2	SENSEI-LIF	$\mathbf{0.630}_{2}$	0.670_{1}	0.617_{7}
3	UNIMELB	$\mathbf{0.617}_{3}$	0.641_{5}	0.616_{8}
4	INESC-ID	$\mathbf{0.610}_{4}$	0.662_{3}	0.600_{10}
5	aueb.twitter.sentiment	$\mathbf{0.605}_{5}$	0.644_{4}	0.629_{6}
6	SentiSys	$\mathbf{0.598}_{6}$	0.641_{5}	0.609_{9}
7	I2RNTU	$\mathbf{0.596}_{7}$	0.637_{7}	0.593_{12}
8	INSIGHT-1	$\mathbf{0.593}_{8}$	0.616_{11}	0.635_{5}
9	TwiSE	$\mathbf{0.586}_{9}$	0.598_{16}	0.528_{24}
10	ECNU (*)	$\mathbf{0.585}_{10}$	0.617_{10}	0.571_{16}
11	NTNUSentEval	$\mathbf{0.583}_{11}$	0.619_{8}	0.643_{2}
12	MDSENT	$\mathbf{0.580}_{12}$	0.592_{18}	0.545_{20}
	CUFE	$\mathbf{0.580}_{12}$	0.619_{8}	0.637_{4}
14	THUIR	$\mathbf{0.576}_{14}$	0.605_{15}	0.596_{11}
	PUT	$\mathbf{0.576}_{14}$	0.607_{13}	0.584_{14}
16	LYS	$\mathbf{0.575}_{16}$	0.615_{12}	0.585_{13}
17	IIP	$\mathbf{0.574}_{17}$	0.579_{19}	0.537_{23}
18	UniPI	$\mathbf{0.571}_{18}$	0.607_{13}	0.639_{3}
19	DIEGOLab16 (*)	$\mathbf{0.554}_{19}$	0.593_{17}	0.549_{19}
20	GTI	$\mathbf{0.539}_{20}$	0.557_{21}	0.518_{26}
21	OPAL	$\mathbf{0.505}_{21}$	0.560_{20}	0.541_{22}
22	DSIC-ELIRF	$\mathbf{0.502}_{22}$	0.511_{25}	0.513_{27}
23	UofL	$\mathbf{0.499}_{23}$	0.537_{22}	0.572_{15}
	ELiRF	$\mathbf{0.499}_{23}$	0.516_{24}	0.543_{21}
25	ISTI-CNR	$\mathbf{0.494}_{25}$	0.529_{23}	0.567_{17}
26	SteM	$\mathbf{0.478}_{26}$	0.496_{27}	0.452_{31}
27	Tweester	$\mathbf{0.455}_{27}$	0.503_{26}	0.523_{25}
28	Minions	$\mathbf{0.415}_{28}$	0.485_{28}	0.556_{18}
29	Aicyber	$\mathbf{0.402}_{29}$	0.457_{29}	0.506_{28}
30	mib	$\mathbf{0.401}_{30}$	0.438_{30}	0.480_{29}
31	VCU-TSA	$\mathbf{0.372}_{31}$	0.390_{32}	0.382_{32}
32	SentimentalITists	$\mathbf{0.339}_{32}$	0.424_{31}	0.480_{29}
33	WR	$\mathbf{0.330}_{33}$	0.333_{34}	0.298_{34}
34	CICBUAPnlp	$\mathbf{0.303}_{34}$	0.377_{33}	0.374_{33}
	Baseline	$\mathbf{0.255}$	0.333	0.342

SwissCheese's method

SwissCheese at SemEval-2016 Task 4: Sentiment Classification Using an Ensemble of Convolutional Neural Networks with Distant Supervision

Jan Deriu*
ETH Zurich
Switzerland
jderiu@student.ethz.ch

Maurice Gonzenbach*
ETH Zurich
Switzerland
mauriceg@student.ethz.ch

Fatih Uzdilli
Zurich University of Applied Sciences
Switzerland
uzdi@zhaw.ch

Aurelien Lucchi
ETH Zurich
Switzerland
alucchi@inf.ethz.ch

Valeria De Luca
ETH Zurich
Switzerland
vdeluca@vision.ee.ethz.ch

Martin Jaggi
ETH Zurich
Switzerland
jaggi@inf.ethz.ch

Steps

- Creation of word embeddings for initialization of the first layer.
- *Distant supervised phrase*, the network weights and word embeddings are trained to capture aspects related to sentiment.
- *Supervised phrase*, where the network is trained on the provided supervised training data.

The architecture of CNNs

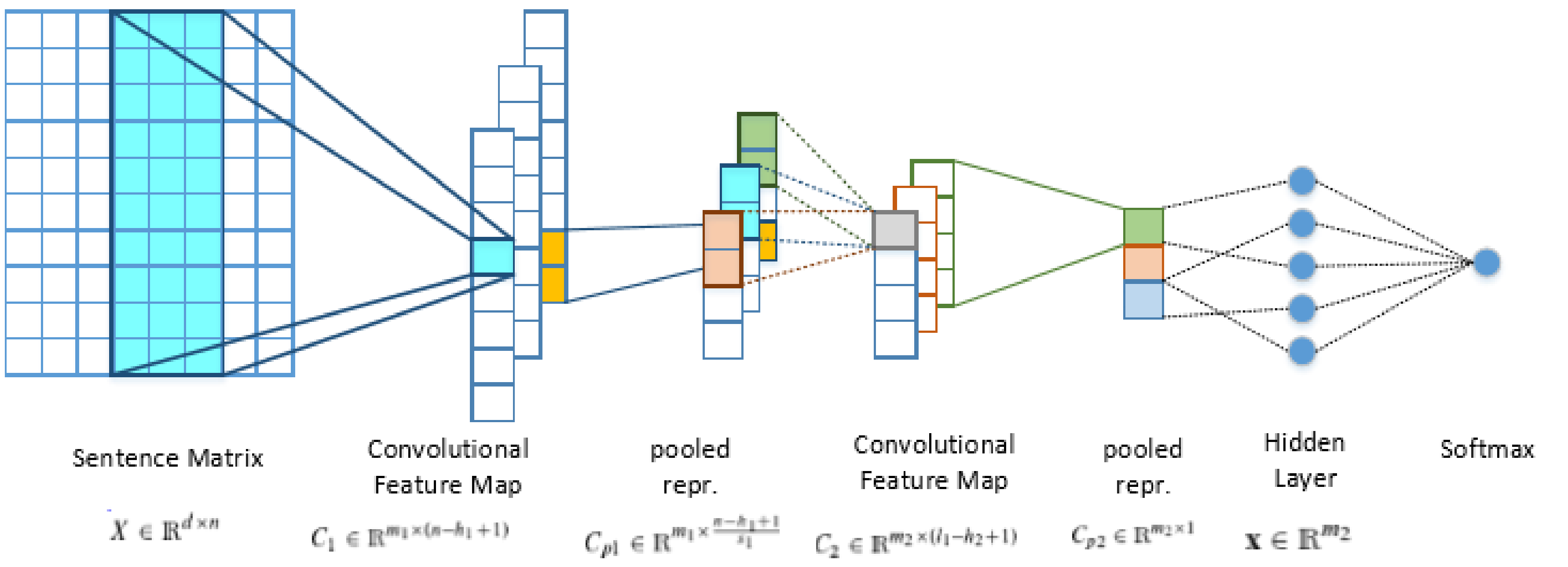


Figure 1: The architecture of the CNNs used in our approach.

Sentence Representation

- A Word → d-dimensional vector
- A Sentence (n constituent words) → d*n dimensional vector

Convolutional Layer

- A set of m filters is applied to a sliding window of length h over a sentence.
- A feature c is generated for a given filter F:

- Aggregated over all m filters into a feature map matrix C with (m * n-h+1)

Max Pooling

- Aggregates vector elements by taking the maximum over a fixed set of non-overlapping invervals

Hidden Layer

- Computes the transformation $\alpha(\mathbf{W} * \mathbf{x} + \mathbf{b})$
- $\mathbf{W} \in \mathbb{R}^{m \times m}$ is the weight matrix, $\mathbf{b} \in \mathbb{R}^{m}$ the bias and alpha the rectified linear (relu) activation function.
- Softmax: return the class:

$$\hat{y} := \arg\max_{j} P(y = j | \mathbf{x}, \mathbf{w}, \mathbf{a})$$
$$= \arg\max_{j} \frac{e^{\mathbf{x}^\intercal \mathbf{w}_j + a_j}}{\sum_{k=1}^{K} e^{\mathbf{x}^\intercal \mathbf{w}_k + a_j}},$$

Ensemble of Classifiers

- Create two 2-layer CNNs

- System 1:

Word vector: Skipgram model of window size 5 and filter words < 5 times, d=52 based on 200M unlabelled twitter corpus.

distant-supervised phase: use emoticons to infer the polarity of a balanced corpus. 90M

Supervised training: 90M distant-supervised labelled, plus labelled data provided.

Ensemble of Classifiers (cont.)

- System 2:

Word vector*: Glove model*, d=50 + 4 different flags (hashtags words, been elongated, all capitalized, punctuations repeated more than 3 times) 90M corpus for POS,NEG and Neutral.

Distant-supervised: 60M balanced corpus.

Supervised training: 60M, plus labelled data provided.

Ensemble of Classifiers

- Meta-classifier
- Input: sentiment class and categorical value of system 1 and II as features
- Model: a random forest using Weka library on the training data.
- 随机森林，指的是利用多棵树对样本进行训练并预测的一种分类器。

Competition in China

Mainly from：计算机学会 and 中文信息学会 Chinese microblog evaluation

Important Information Download:

1. Evaluation Notice
2. 《Registration Form》
3. 《Task guidelines: Emotion Analysis in Chinese Weibo Texts》
4. 《Task guidelines: Sentiment Classification with Deep Learning》
5. 《Task guidelines: Chinese Entity Linking》
6. 《Task guidelines: Cross-Lingual Knowledge Linking》
7. 《Task guidelines: Large Scale English Question Answering》
8. 《Task guidelines: Large Scale Chinese News Categorization》

第八届中文倾向性分析评测(COAE2016)

新增:第八届中文倾向性分析评测（ COAE2016 ）大纲(2016年9月23号更新)

评测大纲下载(2016年9月23号更新)

为了持续推动中文倾向性分析技术的发展和应用， 中文信息学会信息检索专业委员会将在成功 组织前七届中文倾向性分析评测的基础上， 以在华南理工大学举行的第二十二届全国信 息检索学术会议（CCIR2016）为依托，继续组织 第八届中文倾向性分析评测（The 8th Chinese Opinion Analysis Evaluation-COAE2016）。

众所周知，文本倾向性（观点和情感等）分析已经连续多年成为自然语言处 理领域研究的热点问题之一。TREC评测、NTCIR评测以及前七届中文倾 向性分析评测推动和加速了倾向性分析研究的发展。在SIGIR、ACL、WWW、EMNLP、CIKM、WSDM等著名国际会议上，针对这一问题的 研究成果层出不穷。随着研究的深入展开，也出现了一些新的研究关注点， 如Aspect-Based Opinion Mining，Context-sensitive Opinion Mining，Deep Learning based Opinion Mining等。在国内，对于文本倾向性分析的研究蒸 蒸日上。如何结合中文处理的特点，进一步推动中文情感分析的发展是目前亟待 解决的问题。因此，在前七届中文倾向性分析评测的基础上，本届评测将继续致 力于探索中文倾向性分析的新技术、新方法，加强中文倾向性分析理论和技术的研 究及应用， 持续推动建立、完善中文倾向性分析研究的基础资源库和评测标准。

前七届中文倾向性分析评测得到了国内外同行的热情支持和广泛参与，依托CCIR会议举行的

赞助

交通路线

一、 白云机场至华南理

在机场南站乘坐地铁

西路方向），在体育西

号线（番禺广场方向）

5号线（文冲方向），在

4号线（金洲方向）到大

华南理工大学。 more...

SMP 2016 技术评测

简介

参赛队伍利用给定的新浪微博数据（包括用户个人信息、用户微博文本以及用户粉丝列表，详见数据描述部分），进行微博用户画像，具体包括以下三个任务：

任务1：推断用户的年龄（共3个标签：-1979/1980-1989/1990+）

任务2：推断用户的性别（共2个标签：男/女）

任务3：推断用户的地域（共8个标签：东北/华北/华中/华东/西北/西南/华南/境外）

结果揭晓

经过两个多月的激烈角逐，全国社会媒体处理大会首届技术评测竞赛——“微众杯”技术评测（WEIZOOMSMP CUP 2016）于9月30日完成比赛并揭晓结果，来自哈尔滨工业大学深圳研究生院的HLT_HITSZ队斩获冠军（一等奖），另有6支队伍分获二、三等奖。7支获奖队伍将分享2万元奖金，并受邀在10月29至30于江西南昌召开的SMP 2016大会上作技术评测报告并参加颁奖仪式。

Project Option 2 (大作业选项 2)

- Sentiment Analysis (情感分析)
- Input ： reviews of Movies or Products in English and Chinese
- Output：Positive, Negative and Neutral

- 确实非常不错，物有所值哦 （正）
- 很受用啊。有经验的感触更深！ （正）
- 没想到是这种音乐，效果很好，但是歌太老了。（负）

Schedules

- **Before Nov.18:**： submit team task and members

li-fang@cs.sjtu.edu.cn

- **Dec.28**： Each group presents in a workshop.

PPT includes:

task description (indicate each person's subtask)

Method description

Steps to implement

preliminary experiments.

- **Jan,? Morning 9AM~11点AM at ?**： evaluation
- Final submit： **PPT**， **coding**。

Demo of sentiment classification in our research

- 我就说#幸福36计# 好看的[偷乐]   +
- 幸福36计，女主角太他妈的丑了，丑的让我无法直视，第一集没看完直接不看了，丑哭了[泪]   --
- #罗晋# #幸福36计# 不见不散😍   0