

Exploiting emoticons in sentiment analysis

Hirakata

Univ. of Tokyo, IS, B4

20131116

reference

“Exploiting emoticons in sentiment analysis.”

- author: Hogenboom, Alexander, et al.
- Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, 2013.
- <http://eprints.eemcs.utwente.nl/23268/01/sac13-senticon.pdf>

outline

- Intro
- Nature of emoticons
- How to analyze

- analyze sentiment (positive or negative)

- analyze sentiment (positive or negative)
- This paper analyze Twitter tweets and forum messages of Google discussion. That texts include emoticons

- analyze sentiment (positive or negative)
- This paper analyze Twitter tweets and forum messages of Google discussion. That texts include emoticons
- using with lexicon of emoticons and lexicon of words

- analyze sentiment (positive or negative)
- This paper analyze Twitter tweets and forum messages of Google discussion. That texts include emoticons
- using with lexicon of emoticons and lexicon of words
- lexicons are made manually

nature of emoticons

- place of emoticons
- how emoticons are used

place of emoticons

- 84.0 placed at the end of paragraph
- 9.0 placed at the middle of paragraph
- 7.0 placed at the head of paragraph

assume

- an emoticon affects a segment
- segments are sentences or paragraphs

place of emoticons

- 84.0 placed at the end of paragraph
- 9.0 placed at the middle of paragraph
- 7.0 placed at the head of paragraph

assume

- an emoticon affects a segment
- segments are sentences or paragraphs
- if the paragraph has two or more emoticons then sentence
else paragraph

how used

emoticons are used in 3 types

- text not express sentiment clearly
- emoticons stress sentiment
- negation

how used

emoticons are used in 3 types

- text not express sentiment clearly
 - “(^^)I got a promotion.”
- emoticons stress sentiment
- negation

how used

emoticons are used in 3 types

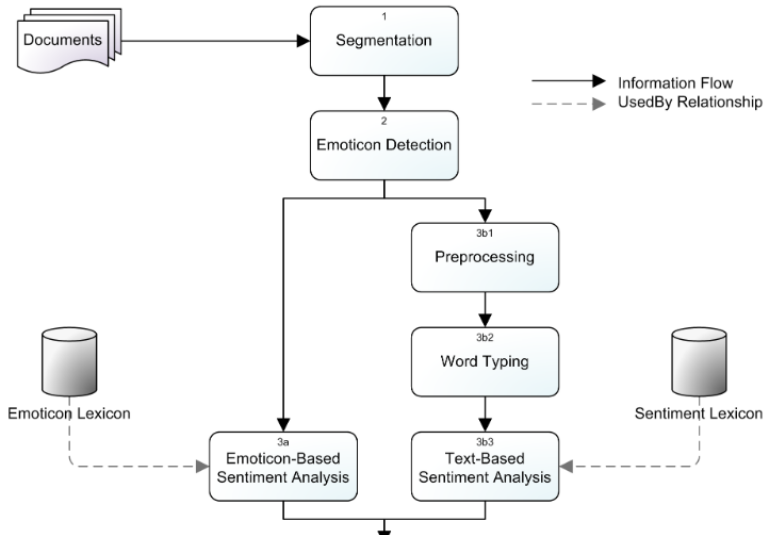
- text not express sentiment clearly
 - “(^^)I got a promotion.”
- emoticons stress sentiment
 - “I love my work (^)”
- negation

how used

emoticons are used in 3 types

- text not express sentiment clearly
 - “(^^)I got a promotion.”
- emoticons stress sentiment
 - “I love my work (^)”
- negation
 - “I love my work (-)”

Overview of analysis [1/2]



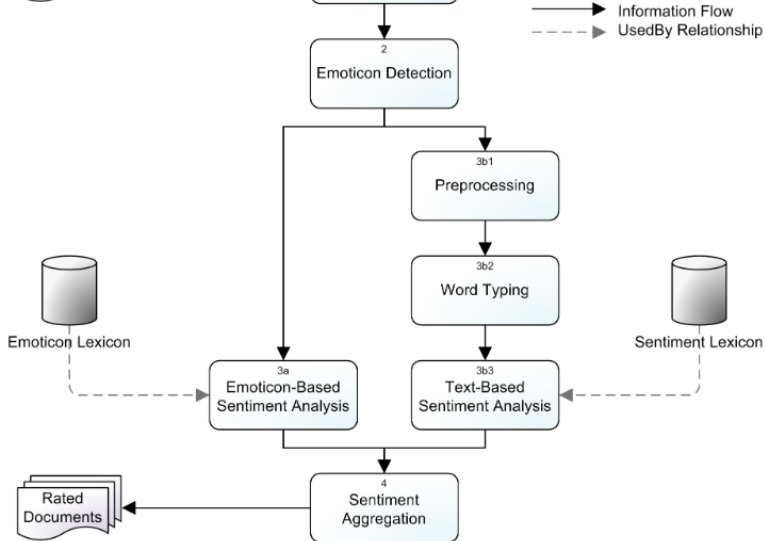


Figure 1: Overview of our sentiment analysis framework.

step 3a — sentiment score of an emoticon

lexicon of emoticons is list of (emoticon \rightarrow score) where score is positive for positive sentiment

emoticon $e \rightarrow \text{sent}(e)$

e.g.

■ $\text{sent}(\text{^^}) = 0.8$

■ $\text{sent}(\text{--}) = -0.7$

(dummy values)

step 3b — sentiment score of a word

sentiment-carrying words and modifyings are used
lexicon of emoticons is list of (word \rightarrow score) where score is positive for positive sentiment

word $w \rightarrow sent_w$

e.g. $sent \text{ "lol" } = 0.5$

step 4 — merge

for when negation emoticons, when emoticons exist, words
(step 3b) are ingored

step 4

for a segment s

- When s has v emoticons,
- i.e. $es = \{e_1 \dots e_v\}$ $sent(es) = \{sent(e_1) \dots sent(e_v)\}$
(notation)

step 4

for a segment s

- When s has v emoticons,
- i.e. $es = \{e_1 \dots e_v\}$ $sent(es) = \{sent(e_1) \dots sent(e_v)\}$
(notation)
- then score of emoticons of a segment s

$$sent_e(s) = \frac{1}{v} \sum sent(es)$$

step 4

for a segment s

- When s has v emoticons,
- i.e. $es = \{e_1 \dots e_v\}$ $sent(es) = \{sent(e_1) \dots sent(e_v)\}$
(notation)
- then score of emoticons of a segment s

$$sent_e(s) = \frac{1}{v} \sum sent(es)$$

- When s has t sentiment-carrying words and modifyings,
- $ws = \{w_1 \dots w_t\}$ (words)
- $ms = \{m_1 \dots m_t\}$ (modifyings)

step 4

for a segment s

- When s has v emoticons,
- i.e. $es = \{e_1 \dots e_v\}$ $sent(es) = \{sent(e_1) \dots sent(e_v)\}$
(notation)
- then score of emoticons of a segment s

$$sent_e(s) = \frac{1}{v} \sum sent(es)$$

- When s has t sentiment-carrying words and modifyings,
- $ws = \{w_1 \dots w_t\}$ (words)
- $ms = \{m_1 \dots m_t\}$ (modifyings)
- score of words of segment s

$$sent_w(s) = \frac{1}{t} \sum sent(ws)$$

step 4

segment score of a segment

Definition

for a segment s ,

$$sent(s) = \begin{cases} \frac{1}{v} \sum sent(es) & v > 0 \\ \frac{1}{t} \sum sent(ws, ms) & \text{otherwise} \end{cases}$$

segment score of a document d

$$d = \{s_1 \dots s_p\}$$

Definition

$$sent(d) = \frac{\sum_{i=1}^p sent(s_i) \cdot \text{if } v_i > 0 \text{ then } v_i \text{ else } t_i}{\sum_{i=1}^p \text{if } v_i > 0 \text{ then } v_i \text{ else } t_i}$$

evaluation

- $class(d) = \text{if } sent(d) > 0 \text{ then } 1 \text{ else } -1$
- human anotation: $\{-1.0, -0.5, 0, 0.5, 1.0\}$

Result

Table 3: Experimental results for all approaches on a set of documents containing emoticons.

Method	Positive			Negative			Overall	
	Precision	Recall	F_1	Precision	Recall	F_1	Accuracy	Macro F_1
Baseline	0.21	0.22	0.22	0.23	0.22	0.22	0.22	0.22
Sentence-level	0.65	0.67	0.66	0.59	0.68	0.63	0.59	0.65
Paragraph-level	0.95	0.93	0.94	0.93	0.95	0.94	0.94	0.94