ELS: A Word-Level Method for Entity-Level Sentiment Analysis

Nikos Engonopoulos Angeliki Lazaridou Georgios Paliouras
Konstantinos Chandrinos
University of Athens, NCSR "Demokritos", i-sieve Technologies Ltd - Greece

International Conference on Web Intelligence, Mining and Semantics
Sogndal, Norway 2011
This work was partially funded by the syn project.

May 25, 2011

The problem

Task: *identify the sentiment expressed towards entities and their features*

MP3 player review

For the money you get good **[quality]**$_1$ and plenty of **[memory]**$_2$, but you also have to cope with a **[UI]**$_3$ that is far from obvious and is controlled by **[buttons]**$_4$ with a very plastic feel to them.

The problem

Task: *identify the sentiment expressed towards entities and their features*

MP3 player review

For the money you get good **[quality]**$_1$ and plenty of **[memory]**$_2$, but you also have to cope with a **[UI]**$_3$ that is far from obvious and is controlled by **[buttons]**$_4$ with a very plastic feel to them.

Our solution: *sequentially model the sentiment flow*

MP3 player review

For the money you get good **[quality]**$_1$ and plenty of **[memory]**$_2$, but you also have to cope with a **[UI]**$_3$ that is far from obvious and is controlled by **[buttons]**$_4$ with a very plastic feel to them.

Issues in entity-level sentiment analysis

- **High localization:** sentiment about entities expressed in sub-sentential level $\rightarrow$ bag-of-words IR approaches inadequate
- **Domain dependence:** different ways of expressing sentiment across domains $\rightarrow$ rule-based methods not robust
- **Evaluation:** task not obvious, even for human annotators $\rightarrow$ hard to establish gold standard for comparison

Previous approaches

Document-level difficult to infer sentiment towards specific entities

Sentence-level sentence classification is not sufficient for identifying sentiment of entities

Previous approaches

Document-level difficult to infer sentiment towards specific entities

Sentence-level sentence classification is not sufficient for identifying sentiment of entities

Entity-level

[Opine]
- retrieve opinion sentences with extraction rules
- identify context-sensitive polar words
- determine polarity using linguistic information

[HuLiu]
- extract subjective sentences
- identify polarity towards entities contained in the extracted sentences

Overview

Sequential modeling of the word-level **sentiment flow**: the sequence of sentiment labels $Y =< y_1, y_2, ..., y_k >$ corresponding to a sequence of words $X =< x_1, x_2, ..., x_k >$ in a text

Motivation

- sentiment changes within a sentence
- sentiment of a word/phrase depends on context and on previously expressed sentiment

Sentiment flow

[For the money you get good quality and plenty of memory,] [but you also have to cope with a UI that is far from obvious and is controlled by buttons with a very plastic feel to them.]

Overview

Sequential modeling of the word-level **sentiment flow**: the sequence of sentiment labels $Y =< y_1, y_2, ..., y_k >$ corresponding to a sequence of words $X =< x_1, x_2, ..., x_k >$ in a text

Motivation

- sentiment changes within a sentence
- sentiment of a word/phrase depends on context and on previously expressed sentiment

Entity references

[For the money you get good **[quality]**$_1$ and plenty of **[memory]**$_2$,] [but you also have to cope with a **[UI]**$_3$ that is far from obvious and is controlled by **[buttons]**$_4$ with a very plastic feel to them.]

Overview

Sequential modeling of the word-level **sentiment flow**: the sequence of sentiment labels $Y =< y_1, y_2, ..., y_k >$ corresponding to a sequence of words $X =< x_1, x_2, ..., x_k >$ in a text

Motivation

- sentiment changes within a sentence
- sentiment of a word/phrase depends on context and on previously expressed sentiment

Entity-level sentiment

For the money you get good **[quality]**$_1$ and plenty of **[memory]**$_2$, but you also have to cope with a **[UI]**$_3$ that is far from obvious and is controlled by **[buttons]**$_4$ with a very plastic feel to them.

Word-level sentiment modeling

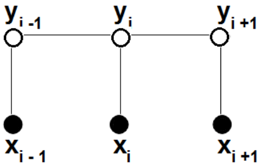
- Training data labeled with:
  1. entity references
  2. segments and their sentiment
- The sentiment label of the segment is passed on to each of its words, creating pairs <word, sentiment>
- Each document is modeled as a sequence of observations (words) and underlying states (sentiment labels)
- Conditional Random Fields (CRF) are used to model this sequence (as implemented in the Mallet toolkit)

Linear-chain Conditional Random Fields

- Discriminative model - scales well to large sets of features
- Dependencies between labels (states), input sequences are learned and weighted through the training data
- Conditional probability is computed as

$$p(Y|X) = \frac{1}{Z(X)} \exp(\sum_{t=1}^{T} \sum_{k=1}^{K} \lambda_k f_k(y_t, y_{t-1}, x_t)) \quad (1)$$

Figure: Example of a linear-chain CRF



Feature vector

- Feature vector of word: word,POS + context of word
- Every document of length k represented as a sequence of k feature vectors

Extract

[...] But/CC at/IN the/DT same/JJ **time/NN** it/PRP takes/VBZ [...]

Table: Feature vector with context window size 5

word_{i-2}	tag_{i-2}	word_{i-1}	tag_{i-1}	word_i	tag_i	word_{i+1}	tag_{i+1}	word_{i+2}	tag_{i+2}
the	DT	same	JJ	time	NN	it	PRP	takes	VBZ

Training: feature vector of word + sentiment label of word

Decoding entity-level sentiment

- Each document is assigned a word-level sequence of sentiment labels

Sentiment flow of document

[...] Creative is an excellent mp3 player, but its supplied earphones are of inferior quality [...]

- **The entity-level sentiment is extracted by the labels assigned to entity references**

Extract local sentiment for entity references

[...] [**Creative**]$_1$ is an excellent mp3 player, but its supplied [**earphones**]$_2$ are of inferior quality [...]

Dataset

- **Dataset:** Customer Review Data [HuLiu]
  - 314 reviews for 5 products
  - 2108 annotated pairs <entity reference, sentiment> (1363 positive, 745 negative, 0 neutral)
- Further annotated with segments and their sentiment
- 72461 annotated words - ∼87% agreement with gold standard on entity level
- Force 100% agreement on entity-level annotation (only pos, neg) for comparison

Entity-level results

After randomly permutating the dataset, we performed a 10-fold cross-validation:

Table: Entity-level sentiment classification

ELS accuracy	H&L opinion recall	H&L polarity accuracy	H&L expected accuracy *
68.6%	69.3%	84.2%	58.4% *

Table: Entity-level opinion recall (binary classification)

ELS method	H&L method
87.8%	69.3%

* combination of opinion extraction recall with polarity classification accuracy

Domain independence experiment

- Aim: test performance on new, unseen types of reviews
- Training set: reviews for 3 of the 4 product types
- Test set: the 4th product type

Table: Domain independence experiment results

	Average for 4 product types	*Initial experiment*
Entity-level accuracy	61.7%	68.6%

Error analysis using pattern discovery

- Frequent patterns observed in the predicted sentiment flow
- Correlation between some word-level prediction sequences and certain types of entity-level error
- Odds ratio: $r = \frac{P(y_{\mathrm{t}} \rightarrow \hat{y}_{\mathrm{f}} | Y)}{P(y_{\mathrm{t}} \rightarrow \hat{y}_{\mathrm{f}})}$
- Significant patterns:
  - positive followed by neutral: decreased probability of error negative→neutral (odds ratio: 0.671)
  - neu-neg-neu: decreased probability of error pos→neu and pos→neg (odds ratio: 0.66, 0.69 resp.)
  - Generally, absence of a label from an alternation pattern in the prediction adds confidence to the absence of a label from the original data - could be used for providing confidence scores

Conclusion

- A method for entity-level sentiment classification using word-level modeling of the sentiment flow
- Advantages:
    - Better performance than previous approaches on entity-level sentiment classification
    - Relatively stable when tested on unseen domains
    - The sentiment flow can be used for error analysis and for detecting higher-level patterns
- Disadvantages:
    - Rich manual annotation needed
- Currently working towards more generic and linguistically-aware approaches needing fewer annotated data

Bibliography

M. Hu and B. Liu.
Mining and summarizing customer reviews.
In Proceedings of ACM Special Interest Group on Knowledge Discovery and Data Mining, pages 168–177, 2004.

A.-M. Popescu and O. Etzioni.
Extracting product features and opinions from reviews.
In Proceedings of Empirical Methods on Natural Language Processing, pages 339–346, 2005.