Pinpointing Sentence-level Subjectivity through Balanced Subjective and Objective Features

Munhyong Kim and Hyopil Shin

{likerainsun, hpshin}@snu.ac.kr

PolTAL 2014

Motivation

- Why Sentence Subjectivity Classification?
  - Recognizing Objective Sentences
    - helps to summarize facts
    - helps to evaluate how objectively a document is written.
      - Wikipedia, newspaper, technical manual
  - Recognizing Subjective Sentences
    - enables collecting only opinions from product reviews and opinions about politicians or issues from SNS.

Previous Studies

- Wiebe et al. (1999)
  - Sentence-level subjectivity classification task (Naïve Bayes)
  - Features: presence or absence of certain syntactic classes, punctuation markers, sentence position in a document
  - Subjectivity (Prec: 69%, Rec: 97%), Objectivity (Prec: 38%, Rec: 3%)
- Wiebe and Rilloff (2005)
  - Subjective clues (from other studies), automatically extracted subjective and objective patterns, and word lists (subjective expressions)
  - Subjectivity (F1: 74.2%), Objectivity (F1: 72.7%)
- No study yet showed why finding subjective and objective sentences is difficult to be balanced and how to achieve more balanced result.

What to Show

To show how to find balanced features for sentence subjectivity classification task, considering its particular characteristics

What is a Subjective Sentence?

- Difficulty in Defining a Subjective Sentence
  - The vagueness of sentence subjectivity

> The conference opened with fierce demonstrations by the opponents of economic globalization…

- Definitions from Other Studies
  - "A sentence is considered to be subjective if it contains one or more private state expressions" (Wilson, 2008)
  - An opinionated sentence includes explicit and implicit opinions
    - The voice quality of this phone is amazing. (explicit)
    - The earphone broke in two days. (implicit) (Liu, 2010)

What is a Subjective Sentence?

- Two Important Properties for Sentence Subjecitivity
  - 1) The source of subjectivity is the author, not any other opinion holders.
  - 2) A sentence can hold both personal opinions and objective facts.
    - The judgment of sentence subjectivity depends on the reader.
- A Proposed Definition of a Subjective Sentence

A sentence is subjective, if the reader perceives the writer's subjectivity towards something from explicit words or implicit contextual information of any kind.

Characteristics of Sentence Subjectivity Classification Task

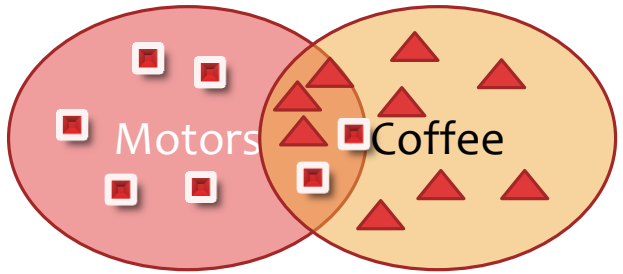
1. The features are asymmetric for subjectivity classification

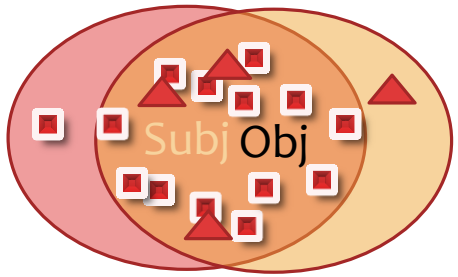
❖ **Usual classification problem**

- ex) document topic classification: motors and coffee
- Each topic has expressions that are likely occur only in that topic.

❖ **Sentence Subjectivity Classification**

- A subjective sentence is likely to have more # of subjective words, but it does not ensure its subjectivity.
- There are no words that occur only in a subjective or objective sentence.





Characteristics of Sentence Subjectivity Classification Task

2. **Sentence Subjectivity is determined not by all subjective words, but a few in the sentence.**
   - main predicates, adverbs, conjunctions, and auxiliary verbs

The Bonn Conference however, which dealt with the global environment and examined the condition of world climate, could be appraised as an exception and a relative success.

Experiment: Task

- Sentence Subjectivity Classification Task
  - Linear SVM model (Vapnik, 1995)
  - Goal: to find balanced features for both subjective and objective sentences

Experiment: Data

- KOSAC: Korean Sentiment Analysis Corpus (Shin et al., 2012, 2013)
  - 7713 sentences (2658 subjective, 5055 objective)
  - 17615 hand-annotated subjective expressions
  - Attributes of each expressions (modified from MPQA corpus schema)
    - polarity, intensity, source, target
    - expressive types: Direct-explicit, Direct-speech, Direct-action, Indirect, and Writing-device
    - semantic types: Judgment, Emotion, Argument, Agreement, Intention, Speculation, Others
- KOSAC DICT
  - automatically extracted subjective expressions from KOSAC

Experiment: Features

- Balanced vs. Imbalanced
  - A feature is imbalanced, if it helps to find only one class, not both.
  - Imbalanced
    - # of Subjective Expressions
      - Subjective sentences do not necessarily contain more # of subjective expressions than objective sentences.
      - High recall for subjective sentence class, low recall for objective sentence.
    - Subjective dictionary features
  - Balanced
    - Probability of a sentence to be subjective
      - balanced recall for subjective and objective class

Experiment: Features

- Expression Frequency Features (SF, N-SF, T-SF)
  - SF: frequency of subjective expressions
  - N-SF: sentence-length based normalized SF
  - T-SF: three-valued (0, 1, or >2) frequency

Experiment: Features

- Probability Features (PES, PSS, PPOSS, PSS-POS groups)
  - PES group: the Probability of an Expression to be Subjective
    - PES = $\frac{Occurrences\ of\ a\ Subjective\ Word\ as\ Subjective}{Total\ Occurrences\ of\ a\ Subjective\ Word\ (N_t)} \times \frac{N_t}{N_t+1}$
    - ex) feature: like/VV-PES, value: 0.653
  - PSS group: the Probability of a Sentence to be Subjective that the expression belongs to
    - PSS = $\frac{\#\ of subjective\ sentences\ that\ the\ Expression\ Belongs\ to}{Total\ \#\ of\ Sentences\ that\ the\ Expression\ Belongs\ to (N_t)} \times \frac{N_t}{N_t+1}$
    - ex) feature: like/VV-PSS, value: 0.43

Experiment: Features

- Probability Features (PES, PSS, PPOSS, PSS-POS groups)
  - PPOSS group: the Probability of Part-Of-Speech sequence Subjectivity
    - ex) feature: VV-PPOSS, value: 0.32
  - PSS-POS group: the Probability of Subjectivity of the Sentence that the Part-Of-Speech sequence belongs to
    - ex) feature: VV-PPOSS, value: 0.12
- Averaged Probability Features (AVG-PES*, AVG-PSS*, AVG-PPOSS*, AVG-PSS-POS*)
  - PES, PSS, PPOSS, PSS-POS of all subjective words are averaged for each sentence.

Experiment: Features

- Intensity Features (CAT-INT, CONT-INT, AVG-CONT-INT*)
  - CAT-INT group: categorical intensity feature
    - enjoy/vv-CAT-INT: medium
  - CONT-INT group: continuous intensity feature
    - enjoy/vv-CONT-INT : 1~3 continuously scaled from all annotations of the expression
  - AVG-CONT-INT*: one averaged continuous intensity value of all subjective expressions in a sentence

Experiment: Features

- Expressive Type Features (EXP-TYPE-FREQ*, EXP-TYPE-NORM*, EXP-TYPE-THREE*)
  - EXP-TYPE-FREQ: the frequency of each expressive type in a sentence
    - Ex) direct-speech: 2, direct-action: 0, indirect: 5, writing-device: 1,
  - EXP-TYPE-NORM: the normalized frequency by its sentence length
  - EXP-TYPE-THREE: 0, 1, >2 valued frequency

Experiment: Features

- Semantic Type Features (SEM-TYPE-FREQ*, SEM-TYPE-NORM*, SEM-TYPE-THREE*)
  - SEM-TYPE-FREQ: the frequency of each semantic type in a sentence
  - SEM-TYPE-NORM: the normalized frequency by its sentence length
  - SEM-TYPE-THREE: 0, 1, >2 valued frequency
- Bag-of-Words Features (BOW)
  - words from subjective or objective sentences respectively

Experiment Result
: Separate Feature Examination

Table 1. Separate Feature Examination within KOSAC Corpus

N.	Feature Combination	Precision (SUBJ / OBJ)	Recall (SUBJ / OBJ)	F_1 (SUBJ / OBJ)	Acc.
(1)	SF	54.2 / 74.1	90.7 / 26.0	67.9 / 38.4	57.7
(2)	N-SF	53.4 / 84.8	96.6 / 18.7	68.8 / 30.6	57.0
(3)	T-SF	53.0 / 80.5	95.5 / 18.0	68.1 / 29.4	56.1
(4)	N-SF + PES	55.8 / 84.4	94.8 / 27.3	70.2 / 41.2	60.4
(5)	N-SF + PSS	56.3 / 81.9	93.1 / 30.2	70.1 / 44.0	61.0
(6)	N-SF + AVG-PES*	57.8 / 78.3	89.4 / 37.0	70.2 / 50.2	62.8
(7)	N-SF + EXP-TYPE-FREQ*	56.5 / 81.6	92.7 / 31.2	70.2 / 45.1	61.4
(8)	N-SF + SEM-TYPE-FREQ*	56.6 / 81.1	92.4 / 31.5	70.1 / 45.3	61.4
(9)	N-SF + CAT-INT	61.3 / 71.9	78.9 / 51.8	68.9 / 60.1	65.1
(10)	N-SF + CONT-INT	61.0 / 72.4	80.1 / 50.7	69.2 / 59.5	65.0

Experiment Result: Feature Combination Examination

Table 2. Feature Combination Examination within KOSAC Corpus[1]

N.	Feature Combination	Precision	Recall	F_1	Acc.
(1)	N-SF	53.44/84.87	96.57/18.71	68.77/30.61	56.94
(2)	N-SF + AVG-PES* + AVG-PPOSS*	57.90/78.26	89.33/37.24	70.23/50.42	62.83
(3)	N-SF + AVG-PSS* + AVG-PSS-POS*	52.72/83.40	96.63/16.28	68.18/27.18	55.73
(4)	N-SF + AVG-PES* + AVG-PPOSS* + CAT-INT	65.10/69.15	70.39/63.55	67.53/66.08	66.89
(5)	(4) + CONT-INT	64.23/68.34	69.99/62.22	66.85/64.96	66.00
(6)	(4) + PSS	64.13/70.00	73.10/60.44	68.23/64.75	66.63
(7)	(4) + EXP-TYPE-FREQ* + SEM-TYPE-FREQ*	61.10/76.55	84.69/47.92	70.93/58.84	65.96
(8)	(2) + EXP-TYPE-FREQ*	64.39/71.93	75.99/59.37	69.66/64.99	67.52
(9)	(2) + SEM-TYPE-FREQ*	66.36/70.13	71.38/64.96	68.71/67.37	68.11
(10)	ALL-FEATURES	57.33/69.03	80.01/69.03	66.67/52.20	60.85

[1] The asterisks (*) indicate balanced features that appeared to help sentence subjectivity classification in the experiment.

Experiment Result: with Bag-of-Words

- Only Bag-of-Words Features (subj/obj)
  - Pr.: 72.39 / 72.27, Rec.: 72.25 / 73.04, F_1: 72.29 / 72.6, Acc.: 72.80 / 72.02
- Feature combinations + Bag-of-Words Features
  - N-SF + AVG-PES* + AVG-PPOSS* + BOW => Acc.: 73.57%
  - N-SF + AVG-PES* + AVG-PPOSS* + SEM-TYPE-FREQ* + BOW => Acc.: 73.15%

Experiment Result: Validation of Feature Combinations

- Data
  - SEJONG: Sejong Part-of-speech Tagged Corpus
    - 900 sentences (500 subjective, 400 objective)
    - to ensure that the feature combinations in KOSAC are still effective.
  - NEG-WIKI: Wikipedia pages about wars and diseases
    - 308 sentences (all objective)
    - Sentences in NEG-WIKI include high number of negative words.
      - assumed to be difficult to predict their objectvities
  - REPORT-SEJONG: reporting sentences from SEJONG
    - 50 reporting sentences (all objective)
    - Sentences in REPORT-SEJONG include high number of subjective words even though the writer wrote it to objectively report.

Experiment Result: Validation of Feature Combinations

Table 3. Feature Combination Test of SEJONG Corpus with Bag-of-words Features

N.	Feature Combination	Precision	Recall	F_1	Acc.
(1)	BOW + N-SF	68.10/76.49	88.22/48.12	76.87/59.08	70.44
(2)	(1) + AVG-PES*	72.57/68.00	76.05/63.91	74.27/65.89	70.67
(3)	(2) + EXP-TYPE-FREQ*	70.00/74.48	85.23/54.14	76.87/62.70	71.44
(4)	(2) + SEM-TYPE-THREE*	73.99/61.38	69.26/68.92	71.55/64.94	67.00
(5)	(2) + EXP-TYPE-FREQ* + SEM-TYPE-THREE*	66.62/69.84	88.82/64.41	76.13/67.01	71.89
(6)	(5) + CAT-INT	67.13/72.08	87.23/55.64	75.87/62.80	70.78
(7)	(5) + CONT-INT	66.67/72.97	88.22/54.14	75.95/62.16	70.78
(8)	(5) + CAT-INT + CONT-INT	67.07/73.17	87.82/52.63	76.06/61.22	70.44
(9)	(5) + PES	64.56/71.74	89.82/57.89	75.13/64.08	71.22
(10)	(5) + PSS	64.37/71.52	89.42/55.39	74.85/62.43	70.44
(11)	(5) + PSS-POS	64.88/73.40	91.82/57.39	76.03/64.42	71.89

Experiment Result: Validation of Feature Combinations

Table 4-1. REPORT-SEJONG Classification Task

N.	Feature Combination	Precision	Recall	F1	Acc.
(1)	BOW* + N-SF + AVG-PES* + AV G-PPOSS* + SEM-TYPE-FREQ*	100%	78%	87%	78%
(2)	BOW* + N-SF + AVG-PES* + AV G-PPOSS* + CAT-INT	100%	74%	85%	74%

Table 4-2. NEG-WIKI Classification Task

N.	Feature Combination	Precision	Recall	F1	Acc.
(1)	BOW* + N-SF + AVG-PES* + AV G-PPOSS* + SEM-TYPE-FREQ*	100%	64%	78%	64%
(2)	BOW* + N-SF + AVG-PES* + AV G-PPOSS* + CAT-INT	100%	76%	86%	76%

Summary

- Balanced features with low feature dimensions can help to find subjective as well as objective sentences.
- These feature combinations seem to be robust even on cases that could be hard to correctly classified.

Thank you for listening