A Joint Segmentation and Classification Framework for Sentiment Analysis

Duyu Tang\textsuperscript{1}, Furu Wei\textsuperscript{2}, Bing Qin\textsuperscript{1}, Li Dong\textsuperscript{3}, Ting Liu\textsuperscript{1}, Ming Zhou\textsuperscript{2}

\textsuperscript{1} Harbin Institute of Technology
\textsuperscript{2} Microsoft Research
\textsuperscript{3} Beihang University
Sentence-level sentiment classification

- Input: A sentence (e.g. tweet)
- Output: Sentiment polarity of the sentence
  - Positive / Negative
Existing Methods: Pipelined

Training Data → Segmentation Algorithm → Segmentation Results → Polarity Label → Classification Algorithm
Existing Methods: Pipelined

- Training Data
- Segmentation Algorithm
  - Segmentation Results
  - Polarity Label
- Classification Algorithm

that is not bad

(that is not bad)

Positive (+1)
Existing Methods: Pipelined

Training Data → Segmentation Results → Classification Algorithm → Polarity Label

that is not bad

Positive (+1)

Polarity inconsistency between a phrase and the words it contains.

\{ \text{not bad} \leftrightarrow \text{bad} \}
\{ \text{a great deal of} \leftrightarrow \text{great} \}
A Joint Model

Training Data → Segmentation Results → Polarity Label

Segmentation Algorithm

Classification Algorithm
A Joint Model

Segmentation => Classification

Segmentation result is the input for classification.

Classification => Segmentation

Classification answer (right or wrong) can indicate the usefulness of a segmentation.
The Inference Process

that is not bad
The Inference Process

Input

Segmentations

that is not bad

that is not bad

that is not bad

that is not bad

that is not bad

CG

The candidate generation model
The Inference Process

The candidate generation model

The segmentation ranking model
The Inference Process

The candidate generation model

The segmentation ranking model
The Inference Process

- **CG**: The candidate generation model
- **SC**: The sentiment classification model
- **SEG**: The segmentation ranking model

The process involves:
- Input: "that is not bad"
- Segmentations: Generated segments
- Polarity: Classification results (+1 or -1)
- Top K: Selected sentiments

The diagram illustrates the flow of text analysis from input to sentiment classification, highlighting the models involved and their outputs.
The Inference Process

Input: that is not bad

Segmentations:
- that is not bad
- that is not bad
- that is not bad
- that is not bad

Rank:
- 2.3
- 1.6

Top K:
- 0.6
- 0.4

Polarity:
- SC
- CG

CG: The candidate generation model
SC: The sentiment classification model
SEG: The segmentation ranking model

Vote: +1
that is not bad

Polarity: +1
The Training Process

Polarity: +1

Input

Segmentations

The candidate generation model

The use of a model
The Training Process

- Input: that is not bad
- Segmentations:
  - that is not bad
  - that is not bad
- Polarity:
  - +1

The candidate generation model (CG)
The sentiment classification model (SC)

The use of a model
The Training Process

The candidate generation model

The sentiment classification model

The segmentation ranking model

The use of a model

The update of a model
The Training Process

**Input**

that is not bad

**Segmentations**

-1

**Polarity**

1

**Update**

SEG

NO

SEG

0.6

SEG

NO

0.4

SEG

YES

2.3

SEG

YES

1.6

**Rank**

The candidate generation model

The sentiment classification model

The segmentation ranking model

The use of a model

The update of a model
The Training Process

The candidate generation model

The sentiment classification model

The segmentation ranking model

The use of a model

The update of a model
• Candidate Generation Model

• Segmentation Ranking Model

• Sentiment Classification Model
Candidate Generation Model

• A Beam-Search Approach
  – **Step 1**: Learn a phrase table from a large corpora with **word2vec**
  – **Step 2**: Get the possible segmentations with beam-search.
Candidate Generation Model

• A Beam-Search Approach
  – **Step 1**: Learn a **phrase table** from a large corpora with **word2vec**
  – Intuition
    • Identify phrases based on the occurrence frequency of unigrams and bigrams
      \[ \text{freq}(w_i, w_j) = \frac{\text{freq}(w_i, w_j) - \delta}{\text{freq}(w_i) \times \text{freq}(w_j)} \]
  – Run 2-4 times over the corpora to get longer phrases containing more words
Candidate Generation Model

• A Beam-Search Approach
  – **Step 2**: Beam-Search

| Index | Sentence | Beam |
|-------|----------|------|
|       | **that** |      |
|       | **is**   |      |
|       | **not**  |      |
|       | **bad**  |      |

is not bad

is not bad
Candidate Generation Model

• A Beam-Search Approach
  – **Step 2:** Beam-Search

| Index |
|-------|

| Sentence | that | is | not | bad |
|----------|------|----|-----|-----|

The bigram “that it” is NOT contained in the phrase table.
Candidate Generation Model

• A Beam-Search Approach
  – Step 2: Beam-Search

The bigram “is not” is contained in the phrase table.
Candidate Generation Model

• A Beam-Search Approach
  – **Step 2**: Beam-Search

  The bigram “not bad” is contained in the phrase table.
Candidate Generation Model

- A Beam-Search Approach
  - **Step 2**: Beam-Search

| Index |
|-------|

| Phrase Table |
|-------------|
| is not bad |

| Sentence |
|----------|
| that | is | not | bad |

| Beam |
|-----|
| that | that | is | that | is | not | that | is | not | bad |

The trigram "**is not bad**" is contained in the phrase table.
• Candidate Generation Model

• Segmentation Ranking Model

• Sentiment Classification Model
Segmentation Ranking Model

\[ y = f(x) \]

- \( x \) → Segmentation
- \( y \) → Score

- \( that \ is \ not \ bad \)
- \( that \ is \ not \ bad \)
- \( that \ is \ not \ bad \)

\[ y = f(x) \]

\[ \begin{align*}
that \ is \ not \ bad & \quad 0.6 \\
that \ is \ not \ bad & \quad 2.3 \\
that \ is \ not \ bad & \quad 1.6
\end{align*} \]
A Log-Linear Function

\[ y = f(x) \]

\[ y = \exp(b + \sum_i W_i \cdot F_i) \]

Segmentation -> Score

Weight Feature
# Segmentation Features

| Feature 1 | Feature Description |
|-----------|---------------------|
| #unit     | the number of basic computation units in the segmentation candidate |
| #unit/#word | the ratio of units’ number in a candidate to the length of original sentence |
| #word – #unit | the difference between sentence length and the number of basic computational units |
| #unit > 2 | the number of basic computation units composed of more than two words |

| Feature 2 | Phrase embedding features learned from SkipGram. |
## Segmentation Features

| Feature 1 | Feature Description |
|-----------|---------------------|
| #unit     | the number of basic computation units in the segmentation candidate |
| #unit/#word | the ratio of units’ number in a candidate to the length of original sentence |
| #word – #unit | the difference between sentence length and the number of basic computational units |
| #unit > 2 | the number of basic computation units composed of more than two words |

| Feature 2 | Phrase embedding features learned from SkipGram. |

![Diagram](image)

**Embedding Layer**

**Layer \( f \)**

- \text{average}
- \text{max}
- \text{min}

**Layer \( f' \)**

- \text{concatenate}
Training Objective

\[ x \rightarrow y = f(x) \rightarrow y \]

Segmentation | Score

that is not bad  \( \rightarrow \)  that | is | not | bad

Positive (+1)

Sentence | Segmentation
Training Objective

Segmentation

\[ x \rightarrow y = f(x) \rightarrow y \]

Score

Positive (+1)

that is not bad

\[ y = f(x) \rightarrow 0.6 \]

\[ y = f(x) \rightarrow 2.3 \]

\[ y = f(x) \rightarrow 1.6 \]
Training Objective

\[ y = f(x) \]

Segmentation

Score

that is not bad

Positive (+1)

Sentence

Segmentation

Segmentation Score

Polarity

\[ y = f(x) \]

\[ \begin{align*}
0.6 & \rightarrow -1 & X \\
2.3 & \rightarrow +1 & \checkmark \\
1.6 & \rightarrow +1 & \checkmark 
\end{align*} \]
Training Objective

\[ y = f(x) \]

**Sentence**

*that is not bad*

Positive (+1)

**Segmentation**

\[ y = f(x) \]

\[ \begin{align*}
0.6 & \rightarrow -1 \\
2.3 & \rightarrow +1 \\
1.6 & \rightarrow +1
\end{align*} \]

**Hit candidates**

Maximize the scores of the **hit candidates**.
Training Objective

\[ y = \exp(b + \sum_i W_i \cdot F_i) \]

\[ \text{loss} = -\sum_{i=1}^{\left| T \right|} \log\left( \frac{\sum_{j \in H_i} y_{ij}}{\sum_{j' \in A_i} y_{ij'}} \right) + \lambda \sum_k W_k^2 \]

that is not bad

Positive (+1)

Segmentation

Score

Sentence

Segmentation Score

Polarity

Maximize the scores of the hit candidates.
• Candidate Generation Model

• Segmentation Ranking Model

• Sentiment Classification Model
Sentiment Classification Model

• A supervised learning framework

- **Training Data**
- **Feature Representation**
  - Dimension 1
  - Dimension 2
  - ....
  - Dimension N
- **Learning Algorithm**
- **Sentiment Classifier**

- **Massive Tweets**

  - Emoticon
  - All-cap
  - Elongated

- **Embedding Learning**
  - Embedding Feature
  - NRC-Canada Feature

1 2 3 4 5
N N+1 N+2 N+K
## Classification Feature

| Feature | Feature Description | Reference |
|---------|---------------------|-----------|
| All-Caps | the number of words with all characters in upper case | (Mohammad et al. from National Research Council Canada) |
| Emoticon | the presence of positive (or negative) emoticons, whether the last unit is emoticon |   |
| Hashtag | the number of hashtag |   |
| Elongated units | the number of basic computational containing elongated words (with one character repeated more than two times), such as *goood* |   |
| Sentiment lexicon | the number of sentiment words, the score of last sentiment words, the total sentiment score and the maximal sentiment score for each lexicon |   |
| Negation | the number of negations as individual units in a segmentation |   |
| Bag-of-Units | an extension of bag-of-word for a segmentation |   |
| Punctuation | the number of contiguous sequences of dot, question mark and exclamation mark. |   |
| Cluster | the presence of units from each of the 1,000 clusters from Twitter NLP tool (Gimpel et al., 2011) |   |

**Feature 2**

Phrase embedding features learned from SkipGram.
Experiment Results

• Polarity Classification of Tweets
  – Positive/Negative
  – Benchmark dataset from SemEval 2013

|     | Positive | Negative | Total |
|-----|----------|----------|-------|
| Train | 2,642    | 994      | 3,636 |
| Dev   | 408      | 219      | 627   |
| Test  | 1,570    | 601      | 2,171 |
Results

• Comparison with classification methods
Results

• Comparison with pipelined methods

**Pipeline 1**: Bag-of-word segmentation
**Pipeline 2**: Segmentation with maximum phrases.
Summary

• We develop a joint model to learn sentiment-specific sentence segmentor for sentiment classification

• Our method yields comparable performance with the state-of-the-art methods on Twitter sentiment classification.
Thanks